

\*\*\*\*\*

WQEEH (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 15 14:18:15 1998; MasPar time 17.49 Seconds  
Tabular output not generated. 739.991 Million cell updates/sec

Title: >US-08-952-741-2  
Description: (1-516) from US08952741.pep  
Perfect Score: 3873  
Sequence: 1 MKLHNRIISVLLTLLAVV.....ADGWNFTVNGGAVSVWVKQ 516

Scoring table: PAM 150  
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot35  
1:swiss1

Statistics: Mean 50.716; Variance 89.551; scale 0.566

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description                       | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|
| 1          | 3368  | 87.0        | 518    | 1  | AMT6_BACS7 GLUCAN 1.4-ALPHA-MALTO | 0.00e+00  |
| 2          | 2679  | 69.2        | 512    | 1  | AMY_BACLI ALPHA-AMYLASE PRECURSO  | 0.00e+00  |
| 3          | 2645  | 68.3        | 549    | 1  | AMY_BACST ALPHA-AMYLASE PRECURSO  | 0.00e+00  |
| 4          | 2559  | 66.1        | 514    | 1  | AMY_BACAM ALPHA-AMYLASE PRECURSO  | 0.00e+00  |
| 5          | 1403  | 36.2        | 494    | 1  | AMY2_SALT CYTOPLASMIC ALPHA-AMYL  | 1.24e-289 |
| 6          | 1360  | 35.1        | 495    | 1  | AMY2_ECOLI CYTOPLASMIC ALPHA-AMYL | 1.98e-279 |
| 7          | 217   | 5.6         | 548    | 1  | GLUCAN 1.4-ALPHA-MALTO            | 2.65e-20  |
| 8          | 218   | 5.6         | 551    | 1  | GLUCAN 1.4-ALPHA-MALTO            | 1.70e-20  |
| 9          | 202   | 5.2         | 528    | 1  | AMY_BACCI ALPHA-AMYLASE PRECURSO  | 1.81e-17  |
| 10         | 192   | 5.0         | 713    | 1  | RAW-STARCH-DIGESTING A            | 1.30e-15  |
| 11         | 189   | 4.9         | 712    | 1  | CDGT_BACS3 CYCLOMALTODEXTRIN GLUC | 4.62e-15  |
| 12         | 189   | 4.9         | 718    | 1  | CDGT_BACSS CYCLOMALTODEXTRIN GLUC | 4.62e-15  |
| 13         | 185   | 4.8         | 713    | 1  | CDGU_BACCI CYCLOMALTODEXTRIN GLUC | 2.48e-14  |
| 14         | 186   | 4.8         | 718    | 1  | CDGT_BACCI CYCLOMALTODEXTRIN GLUC | 1.63e-14  |
| 15         | 182   | 4.7         | 711    | 1  | CDGT_BACST CYCLOMALTODEXTRIN GLUC | 8.67e-14  |
| 16         | 183   | 4.7         | 713    | 1  | CDGT_BACS0 CYCLOMALTODEXTRIN GLUC | 5.71e-14  |
| 17         | 183   | 4.7         | 713    | 1  | CDGT_BACSP CYCLOMALTODEXTRIN GLUC | 5.71e-14  |
| 18         | 181   | 4.7         | 718    | 1  | CDGT_BACLI CYCLOMALTODEXTRIN GLUC | 1.31e-13  |
| 19         | 172   | 4.4         | 413    | 1  | AMY3_WHEAT ALPHA-AMYLASE AMY3 PRE | 5.32e-12  |
| 20         | 158   | 4.1         | 421    | 1  | AMYA_VIGMU ALPHA-AMYLASE PRECURSO | 1.45e-09  |
| 21         | 147   | 3.8         | 713    | 1  | CDG2_BACMA CYCLOMALTODEXTRIN GLUC | 1.03e-07  |
| 22         | 147   | 3.8         | 919    | 1  | AMY_STRLI ALPHA-AMYLASE PRECURSO  | 1.03e-07  |
| 23         | 145   | 3.7         | 714    | 1  | CDGI_BACMA CYCLOMALTODEXTRIN GLUC | 2.20e-07  |

|    |     |     |      |   |                                   |          |
|----|-----|-----|------|---|-----------------------------------|----------|
| 24 | 138 | 3.6 | 428  | 1 | AMY1_ORISA ALPHA-AMYLASE PRECURSO | 3.02e-06 |
| 25 | 140 | 3.6 | 440  | 1 | AMY3_ORISA ALPHA-AMYLASE ISOZYME  | 1.44e-06 |
| 26 | 140 | 3.6 | 581  | 1 | AMY1_SCHPO PROBABLE ALPHA-AMYLASE | 1.44e-06 |
| 27 | 139 | 3.6 | 710  | 1 | AMY_THETU ALPHA-AMYLASE PRECURSO  | 2.08e-06 |
| 28 | 138 | 3.6 | 1196 | 1 | AMYB_BACPO BETA-AMYLASE (EC 3.2.1 | 3.02e-06 |
| 29 | 135 | 3.5 | 135  | 1 | AMY5_HORVU ALPHA-AMYLASE TYPE B I | 9.08e-06 |
| 30 | 137 | 3.5 | 368  | 1 | AMY3_HORVU ALPHA-AMYLASE TYPE B I | 4.36e-06 |
| 31 | 137 | 3.5 | 427  | 1 | AMY2_HORVU ALPHA-AMYLASE TYPE B I | 4.36e-06 |
| 32 | 135 | 3.5 | 435  | 1 | AMY3D_ORISA ALPHA-AMYLASE ISOZYME | 9.08e-06 |
| 33 | 136 | 3.5 | 443  | 1 | AMY2_ORISA ALPHA-AMYLASE ISOZYME  | 6.30e-06 |
| 34 | 133 | 3.5 | 445  | 1 | AMY2_ORISA ALPHA-AMYLASE ISOZYME  | 6.30e-06 |
| 35 | 136 | 3.4 | 429  | 1 | AMY6_HORVU ALPHA-AMYLASE TYPE B I | 1.88e-05 |
| 36 | 132 | 3.4 | 438  | 1 | AMY1_HORVU ALPHA-AMYLASE TYPE A I | 2.70e-05 |
| 37 | 126 | 3.3 | 153  | 1 | AMY4_HORVU ALPHA-AMYLASE TYPE B I | 2.28e-04 |
| 38 | 127 | 3.3 | 498  | 1 | AMYA_ASPOW ALPHA-AMYLASE A PRECUR | 1.60e-04 |
| 39 | 127 | 3.3 | 499  | 1 | AMYA_ASPOW ALPHA-AMYLASE A PRECUR | 1.60e-04 |
| 40 | 127 | 3.3 | 499  | 1 | AMYB_ASPOW ALPHA-AMYLASE B PRECUR | 1.60e-04 |
| 41 | 129 | 3.3 | 499  | 1 | AMY_ASPSH ALPHA-AMYLASE PRECURSO  | 1.60e-04 |
| 42 | 129 | 3.3 | 717  | 1 | AMYM_BACST MALTOGENIC ALPHA-AMYL  | 7.89e-05 |
| 43 | 128 | 3.3 | 1592 | 1 | GTFL_STRDO GLUCOSYLTRANSFERASE-1  | 1.13e-04 |
| 44 | 128 | 3.3 | 1597 | 1 | GTFL_STRDO GLUCOSYLTRANSFERASE-1  | 1.13e-04 |
| 45 | 124 | 3.2 | 437  | 1 | AMY3E_ORISA ALPHA-AMYLASE ISOZYME | 4.58e-04 |

ALIGNMENTS

|  |  |  |         |
|--|--|--|---------|
| RESULT 1   | STANDARD;  | PRT;   | 518 AA. |
| AD   | AMT6_BACS7   |  |         |
| AC   | PI9571:  |  |         |
| DT   | 01-FEB-1991 (REL. 17, CREATED)   |  |         |
| DT   | 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)                            |  |         |
| DT   | 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)                          |  |         |
| DE   | GLUCAN 1.4-ALPHA-MALTOHEXAOSIDASE PRECURSOR (EC 3.2.1.98) (G6-AMYLASE) |  |         |
| DE   | (MALTOHEXAOSIDE-PRODUCING AMYLASE) (EXO-MALTOHEXAOSIDOLASE).           |  |         |
| OS   | BACILLUS SP. (STRAIN 707).   |  |         |
| OC   | PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE. |  |         |
| RN   | [1]  |  |         |
| RP   | SEQUENCE FROM N.A., AND SEQUENCE OF 34-36.                             |  |         |
| RX   | MEDLINE; 88162814.   |  |         |
| RA   | TSUKAMOTO A., KIMURA K., ISHII Y., TAKANO T., YAMANE K.;               |  |         |
| RL   | BIOCHEM. BIOPHYS. RES. COMMUN. 151:25-31(1988).                        |  |         |
| CC   | - - CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-ALPHA-D-GLUCOSIDIC LINKAGES  |  |         |
| CC   | IN AMYLACEOUS POLYSACCHARIDES SO AS TO REMOVE SUCCESSIVE               |  |         |
| CC   | MALTOHEXAOSIDE RESIDUES FROM THE NON-REDUCING CHAIN ENDS.              |  |         |
| CC   | - - PATHWAY: DEGRADATION OF STARCH.                                    |  |         |
| CC   | - - SUBCELLULAR LOCATION: EXTRACELLULAR.                               |  |         |
| CC   | - - SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO      |  |         |
| CC   | KNOWN AS THE ALPHA-AMYLASE FAMILY.                                     |  |         |
| DR   | EMBL; M18862; G142497; -   |  |         |
| DR   | PIR; A27705; A27705.   |  |         |
| KW   | HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL.               |  |         |
| FT   | SIGNAL 1 33  |  |         |
| FT   | CHAIN 34 518 MALTOHEXAOSIDE-PRODUCING AMYLASE.                         |  |         |
| SQ   | SEQUENCE 518 AA; 59009 MW; FB2B65AE CRC32;                             |  |         |
| Query Match 87.0%; Score 3368; DB 1; Length 518;               |  |  |         |
| Best Local Similarity 82.8%; Pred. No. 0.00e+00;               |  |  |         |
| Matches 427; Conservative 57; Mismatches 30; Indels 2; Gaps 2; |  |  |         |
| Db   | 3  | MRTCKKFLSILLAFLLVITSI-PTFLVDVEAHNGTNGTMMQYFEWYLPDNGNHNRLN  | 61      |
| Qy   | 1  | MKLNRI-IISVLLTLLAVAVLFPYTPAQAHNGTNGTMMQYFEWHLPDNGNHNRLR    | 59      |
| Db   | 62   | SDAENLKSIGITAVWIPPAWKGASONDYGYGAYDLYDGEFNOKGVTRKYGTRSOLOQA | 121     |
| Qy   | 60   | DDAENLKSIGITAVWIPPAWKGTSQNDYGYGAYDLYDGEFNOKGVTRKYGTRSOLOQA | 119     |
| Db   | 122  | VTSLKNGNIQYGVVNMHKGADATAMVRAVEVNPNNRNQEVGTETTEATRFDFPGR    | 181     |
| Qy   | 120  | VTSLKNGNIQYGVVNMHKGADGTGMVNAVEVNSNRNQEISGEYTTAEATWKDFPGR   | 179     |
| Db   | 182  | GNTHSSFKRWYHFDGVDWDOSRRRLNIYFRGHGKAWDWEVDTEGNWYQYLMYADTD   | 241     |

QY 180 GNTHSNFKRWTHFDGTDQSDQSLQNKIYKFGTGKAWDEVDIENGNDYDLYMAYADIM 239  
 Db 242 DHPVVELNRLNNGVWYNTNLGLDGFRIIDAVKHKISYSTRDWINHRSATGKNWFAVAFW 301  
 QY 240 DHPVVELNRLNNGVWYNTNLGLDGFRIIDAVKHKISYSTRDWINHRSATGKNWFAVAFW 299  
 Db 302 KNDLGAENYLOKTNHNSVDFVPLHYNLNASKSGNDYMRNIFNGTGVVORHPSHAYTF 361  
 QY 300 KNDLGAENYLOKTNHNSVDFVPLHYNLNASKSGNDYMRNIFNGTGVVORHPSHAYTF 359  
 Db 362 VDNHDSQPEALESEFVEENFKPLAYALTLTREGQPSVFGYDYGIPTHGVPAMRSKIDP 421  
 QY 360 VDNHDSQPEALESEFVQSWFKPLAYALTLTREGQPSVFGYDYGIPTHGVPAMRSKIDP 419  
 Db 422 ILKARQYAYGKNDYLDHNTIIGTREGTAHPNSGLATIMSDGAGSKWFMVGRNKG 481  
 QY 420 LQARQYAYGQHDYFDHDIIGTREGDSHPNSGLATIMSDGAGSKWFMVGRNKG 479  
 Db 482 QVNSDITNRTGTVINADGNFSGVSGSIWVN 517  
 QY 480 QVNRDITNRSCTVTINADGNFSGVSGSIWVN 515  
 RESULT 2  
 ID AMY\_BACLI STANDARD; PRT; 512 AA.  
 AC P06278;  
 DT 01-JAN-1988 (REL. 06, CREATED)  
 DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1.4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 GN AMYS OR AMYL.  
 OS BACILLUS LICHENIFORMIS.  
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN-ATCC 27811;  
 RX MEDLINE; 86111694.  
 RA YUKU T., NOMURA T., TEZUKA H., TSUBOI A., YAMAGATA H.,  
 RA TSUKAGOSHI N., UDARA S.;  
 RL J. BIOCHEM. 98:1147-1156(1985).  
 RN [2] SEQUENCE FROM N.A.  
 RP MEDLINE; 86195857.  
 RX GRAY G.L., MAINZER S.E., REY M.W., LAMSA M.H., KINDLE K.L.,  
 RA CARMONA C., REQUAD C.;  
 RL J. BACTERIOL. 166:635-643(1986).  
 RN [3] SEQUENCE OF 1-104 FROM N.A.  
 RP MEDLINE; 84185455.  
 RX STEPHENS M.A., ORTLEPP S.A., OLLINGTON J.F., MCCONNELL D.J.;  
 RL J. BACTERIOL. 158:369-372(1984).  
 RN [4] SEQUENCE OF 1-29 FROM N.A.  
 RP MEDLINE; 89213924.  
 RX LAOIDE B.M., CHAMBLISS G.H., MCCONNELL D.J.;  
 RL J. BACTERIOL. 171:2435-2442(1989).  
 RN [5] SEQUENCE OF 30-47.  
 RP MEDLINE; 82098050.  
 RX KUHN H., FIETZKE P.P., LAMPE J.O.;  
 RL J. BACTERIOL. 149:372-373(1982).  
 RN [6] X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RP STRAIN-ATCC 27811;  
 RX MEDLINE; 95182462.  
 RA MACHUS M., WIEGAND G., HUBER R.;  
 RL J. MOL. BIOL. 246:545-559(1995).  
 CC -1- CATALYTIC ACTIVITY: ENDOLYSIN OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 DR EMBL; X03236; G39552; -

DR EMBL; M38570; G142480; -  
 DR EMBL; M13256; G142511; -  
 DR EMBL; K01984; G142433; -  
 DR EMBL; M26412; G516590; -  
 DR EMBL; A17930; G512528; -  
 DR PIR; A00844; ALBSL.  
 DR PIR; B24549; B24549.  
 DR PIR; A26151; A26151.  
 DR PDB; 1BPL; 17-AUG-96.  
 DR PDB; 1VJS; 12-MAR-97.  
 KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL; 3D-STRUCTURE.  
 FT CHAIN 1 29  
 FT SIGNAL 30 512  
 FT ACT\_SITE 260 260  
 FT ACT\_SITE 264 284  
 FT ACT\_SITE 357 357  
 FT CONFLICT 38 38  
 FT CONFLICT 163 163  
 FT CONFLICT 339 339  
 FT CONFLICT 349 349  
 SQ SEQUENCE 512 AA; 58549 MW; 4B5D334D CRC32;  
 Query Match 69.2%; Score 2679; DB 1; Length 512;  
 Best Local Similarity 66.9%; Pred. No. 0.00e+00;  
 Matches 347; Conservative 86; Mismatches 76; Indels 10; Gaps 7;  
 Db 1 MKQOKRLYARLLTLFALIFLPH-SAAAAA--N-NGTLMQYFEMYPNDQOHKRLQN 56  
 QY 1 MKNLRIISVLLTLLAVAFPPYMTPEPAQHNGTNGTMMQYFEWHLPGDNGHNRURD 60  
 Db 57 DSAYLAEGITAVIWPAYKGTSDQVGYGAYDLYDLGFEHOKGTVRKYTKGELQSAI 116  
 QY 61 DAANLASKGITAVIWPAYKGTSDQVGYGAYDLYDLGFEHOKGTVRKYTKGELQSAI 116  
 Db 117 KSLHSRDINVGVDVINHKGGADATEDVAVEPDADNRVISEGHRKAWTHFFPGRG 120  
 QY 121 TSLKNGIOVYGVVYNNHKGADGTEMVAVENRNSRNOEISEGXTIEAWTKFDFPGRG 176  
 Db 177 TSYDFKWHYHFDGTDWDSRKL-NRIYFQG--KAMDWEVSNENGVYDLYMAYADIM 180  
 QY 181 NTHSNFKRWTHFDGTDQSDQSLQNKIYKFGTGKAWDEVDIENGNDYDLYMAYADIM 233  
 Db 234 HPDVAEIKRWGTWYANLQIDGFRIDAVKHKISYSTRDWINHRSATGKNWFAVAFW 240  
 QY 241 HPEVINELNNGVWYNTNLGLDGFRIIDAVKHKISYSTRDWINHRSATGKNWFAVAFW 293  
 Db 294 NDLGAENYLOKTNHNSVDFVPLHYNLNASKSGNDYMRNIFNGTGVVORHPSHAYTF 300  
 QY 301 NDLAENYLOKTNHNSVDFVPLHYNLNASKSGNDYMRNIFNGTGVVORHPSHAYTF 353  
 Db 354 DNHDTPQGSLESTVQTFKPLAYAFILTRGSGYQVFGYDYGIPTHGVPAMRSKIDP 360  
 QY 361 DNHDTPQGSLESTVQTFKPLAYAFILTRGSGYQVFGYDYGIPTHGVPAMRSKIDP 413  
 Db 414 EPILKARKOYAGQHDYFDHDIIGTREGDSHPNSGLATIMSDGAGSKWFMVGRNKG 417  
 QY 418 DPLQARQYAYGQHDYFDHDIIGTREGDSHPNSGLATIMSDGAGSKWFMVGRNKG 473  
 Db 474 AGETHDITGNRSEPVVINSGEHFGVSGSIWVN 512  
 QY 478 AGQVNRDITNRSCTVTINADGNFSGVSGSIWVN 515  
 RESULT 3  
 ID AMY\_BACST STANDARD; PRT; 549 AA.  
 AC P06279;  
 DT 01-JAN-1988 (REL. 06, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1.4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 GN AMYS.  
 OS BACILLUS STEAROTHERMOPHILUS.

OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 35-39.  
 RX MEDLINE; 85234394.  
 RA NAKAJIMA R., IMANAKA T., AIBA S.;  
 RL J. BACTERIOL. 163:401-406(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN-DY5/PHI300;  
 RX MEDLINE; 86008166.  
 RA IHARA H., SASAKI T., TSUBOI A., YAMAGATA H., TSUKAGOSHI N., UDAKA S.;  
 RL J. BIOCHEM. 98:95-103(1985).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-N2-3;  
 RX MEDLINE; 86195857.  
 RA GRAY G.L., MAINZER S.E., REY M.W., LAMSA M.H., KINDLE K.L.,  
 RA CARMONA C., REQUADT C.;  
 RL J. BACTERIOL. 166:635-643(1986).  
 RN [4]  
 RP SEQUENCE OF 1-122 FROM N.A., AND SEQUENCE OF 35-48.  
 RC STRAIN-DY-5;  
 RX MEDLINE; 86059211.  
 RA TSUKAGOSHI N., IRIANI S., SASAKI T., TAKEMURA T., IHARA H.,  
 RA IDOTA Y., YAMAGATA H., UDAKA S.;  
 RL J. BACTERIOL. 184:1182-1187(1985).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 DR EMBL; M11450; G142505; -;  
 DR EMBL; X02769; G580825; -;  
 DR EMBL; M13255; G142513; -;  
 DR PIR; A00845; ALBSF.  
 DR PIR; A24549; A24549.  
 KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL.  
 FT SIGNAL 1 34  
 FT CHAIN 35 549  
 FT ACT\_SITE 268 268  
 FT ACT\_SITE 272 272  
 FT ACT\_SITE 365 365  
 FT CONFLICT 13 13  
 FT CONFLICT 19 19  
 FT CONFLICT 23 23  
 FT CONFLICT 31 31  
 FT CONFLICT 107 107  
 FT CONFLICT 179 179  
 FT CONFLICT 251 251  
 FT CONFLICT 284 284  
 FT CONFLICT 312 312  
 FT CONFLICT 338 338  
 FT CONFLICT 342 342  
 FT CONFLICT 346 346  
 FT CONFLICT 376 376  
 FT CONFLICT 526 526  
 FT CONFLICT 527 527  
 FT CONFLICT 535 535  
 SQ SEQUENCE 549 AA; 62670 MW; 4C7BEAD6 CRC32;  
 Query Match 68.3%; Score 2645; DB 1; Length 549;  
 Best Local Similarity 67.1%; Pred. No. 0.00e+00;  
 Matches 339; Conservative 81; Mismatches 80; Indels 5; Gaps 5;  
 Db 15 LLAFLT-ALLFCPTGQAKAA-AFENGTMQYFEMWLPDDGLTWKTVANEANLSSIGI 72  
 Qy 11 LUTLLAVAVLPYPYTEPAQAHHNTGNTGMQYFHWLHPNDGNHWNRLRDAANLKSIGI 70  
 Db 73 TALMLPPAYKTSRSDVGYGYDYLDFEENQKQAVRTKYGTKAQYLO-AIOAAHAACMQ 131  
 Qy 71 TAVWIPPAWKTSQNDVGYGYDYLDFEENQKQAVRTKYGTGRSQ-LOGAVTSLKNNGIQ 129  
 Db 132 VYADVVFHKGAGDTEWDAVEVNPDRNQEISTYIOQAWTKFDFPGRGNTYSSFKWR 191  
 FT SIGNAL 1 31  
 FT CHAIN 32 514

QY 130 VYGVVANHKGAGDTEWDAVEVNPDRNQEISTYIOQAWTKFDFPGRGNTYSSFKWR 189  
 Db 192 WYHFDGYDWDSESKL-SRIYKFRGIGKAWDEVDTEGNTDYLMYADLMDHDPVVTTELK 250  
 QY 190 WYHFDGYDWDSESKL-SRIYKFRGIGKAWDEVDTEGNTDYLMYADLMDHDPVVTTELK 249  
 Db 251 SWGKVVNTNIDGRLDAVKHKKFSPDWSVRSQTKPLFTVGEYSYDINKLHNY 310  
 QY 250 NWGVYNTNLNLDGRLDAVKHKKFSPDWSVRSQTKPLFTVGEYSYDINKLHNY 309  
 Db 311 IMKTNGTMSLFDAPLHNKFTASKSGGTFDMRTLMTLTKMDQPTLAVTFVDNHDTEPGQ 370  
 QY 310 LNKTSWNHSEVDFPLHNLNLSNNGGTFDMRTLMTLTKMDQPTLAVTFVDNHDTEPGQ 369  
 Db 371 ALQSWDVPWFKPLAYALITRQEGYPCVFGYGYQYINIPISLKSIDPLLIARRDYAY 430  
 QY 370 ALESFVQSWFPLAYALITRQEGYPCVFGYGYQYINIPISLKSIDPLLIARRDYAY 429  
 Db 431 GTQHDYLDHSDIIGWTRGVTEKPGSLAALITDGPQSGKMWYGVKHAGKVFDLTGNR 490  
 QY 430 GTQHDYLDHSDIIGWTRGVTEKPGSLAALITDGPQSGKMWYGVKHAGKVFDLTGNR 489  
 Db 491 SDVTYNSDGMGFEKVGSGVSWV 515  
 QY 490 SGTVTINADGMGNETVNGGAVSWV 514  
 RESULT 4  
 ID AMY BACAM STANDARD; PRT; 514 AA.  
 AC P00692;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 OS BACILLUS AMYLOLIQUEFACIENS.  
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-IH;  
 RX MEDLINE; 83108808.  
 RA TAKKINEN K., PETERSSON R.F., KALKKINEN N., PALVA I., SODERLUND H.,  
 RA KAARTAINEN L.;  
 RL J. BIOL. CHEM. 258:1007-1013(1983).  
 RN [2]  
 RP SEQUENCE OF 32-222.  
 RX MEDLINE; 80241725.  
 RA CHUNG H.S., FRIEDBERG F.;  
 RL BIOCHEM. J. 185:387-395(1980).  
 RN [3]  
 RP SEQUENCE OF 1-96 FROM N.A.  
 RX MEDLINE; 82051296.  
 RA PALVA I., PETERSSON R.F., KALKKINEN N., LEHTOVAARA P., SARVAS M.,  
 RA SODERLUND H., TAKKINEN K., KAARTAINEN L.;  
 RL GENE 15:43-51(1981).  
 RN [4]  
 RP SEQUENCE OF 1-39 FROM N.A.  
 RX MEDLINE; 88137952.  
 RA RUOHONEN L., HACKMAN P., LEHTOVAARA P., KNOWLES J.K.C., KARAENEN S.;  
 RL GENE 59:161-170(1987).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 DR EMBL; J01542; G142429; -;  
 DR EMBL; V00092; G39298; -;  
 DR EMBL; A20154; G580683; -;  
 DR EMBL; M18424; G142431; -;  
 DR PIR; A00843; ALBSN.  
 KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL.  
 FT SIGNAL 1 31  
 FT CHAIN 32 514

US-08-952-741-2.ISP

Wed Sep 16 09:49:09 1998

J. GEN. MICROBIOL. 138:1051-1065(1992).

[3]  
 RN SEQUENCE OF 476-494 FROM N.A.  
 RP RAHA M., KIHARA M., KAWAGISHI I., MACNAB R.M.;  
 RA SUBMITTED (MAR-1993) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RL CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 DR EMBL: L01643; G154045; -  
 DR EMBL: M85241; G153989; -  
 DR EMBL: L13280; G295194; -  
 DR PIR: B45738; B45738.  
 DR STYGENE; SG10011; AMYA.  
 KW HYDROLASE; GLUCOSIDASE; CARBOHYDRATE METABOLISM.  
 FT ACT\_SITE 235 235 BY SIMILARITY.  
 FT ACT\_SITE 239 239 BY SIMILARITY.  
 FT ACT\_SITE 332 332 BY SIMILARITY.  
 FT ACT\_SITE 494 AA; 56496 MW; 2F27D4B3 CRC32;  
 SQ SEQUENCE 494 AA; 56496 MW; 2F27D4B3 CRC32;  
 Query Match 36.2%; Score 1403; DB 1; Length 494;  
 Best Local Similarity 42.5%; Pred. No. 1.24e-289;  
 Matches 209; Conservative 106; Mismatches 159; Indels 18; Gaps 15;

Db 3 NPTLQYFHWYDPDGKLSWELAEADGLNDGINMVLPPACKGASGGYSGYDYDLF 62  
 QY 37 NGTMQYFEWHLNDGNHNRRLRDAANLKSIGITAVWIPPAWKTSGN-DVGYGAYDLY 95  
 Db 63 DLGEFQKGTIATKYDKRQLLTALDALKNNIAVLDDVYVNHKMGADKERIRVQRNQ 122  
 QY 96 DLGEFQKGTIRYKTSRQLOGAVTSLKNGIOVYGVNMHKGADGTEMVNAVEVR 155  
 Db 123 DDRQTDNIIIECEGTRTYFPARAGQYSNFTWD-YHCFSGIDHINPD-EDGIFKIYND 180  
 QY 156 SNRNOEISGEYITEAMTKDFPGRGTHSNFKRWYH-EDGTDMDQSRQLQNKYKF-RG 213  
 Db 181 YTGQWMDQVDDMGNFYDLMLGENIDFRNHVATEEIKYKARVMEQTHCDGFRDVAVKHI 240  
 QY 214 TGAWDEVDIENGNDYLMYADIDMDHPEVINELRNMGWYITNLNLDGFRDVAVKHI 272  
 Db 241 PAFWKEWIEHVQAVAKPLFIYAEVSWHEVDKLOTYYIDVDGKTMFDAPLQMFHEAS 300  
 QY 273 KYSTRDLTHVRNWTGKPMFAEAFKNDLAAIENLNKTSWNHVSFVDPVLYLYNAS 332  
 Db 301 RGAEDYDMRHFTGTLEADPFHATLVVAHNDTQPOALEAPVPEWFKPLAYALLILREN 360  
 QY 333 NSGGYFDMRNLNGSVQKHPHATVFDNHDSPGSELESFVQSWFKPLAYALLILREQ 392  
 Db 361 GVPVSFYPOLYGASVEDSGENGETCRVDMVPYINOLDRLILARQFAHGQTLEFDHPNCI 420  
 QY 393 GYPSVEFYGD-Y--YGIPT-HG-V-P-SMK--SKIDPLQARQTYAYGTQHDYFDHDDII 442  
 Db 421 AFSRSGTEENPGC-VV-VLSNGDDGKTLILLGDNYANKTRWDFSGNDEYVVTNDQGEAT 478  
 QY 443 GWTRGESSHPNSGLATINSDPGGNKMYVGHKKGAGQWRDITGNRSGTIVNADGWN 502  
 Db 479 FFCNAGSVSVWV 490  
 QY 503 FTVNGGAVSVWV 514

RESULT 6  
 ID AMY2.ECOLI STANDARD; PRT; 495 AA.  
 AC P26612; P78072;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE CYTOPLASMIC ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 GN AMYA.  
 OS ESCHERICHIA COLI.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;

FT ACT\_SITE 262 262 BY SIMILARITY.  
 FT ACT\_SITE 266 266 BY SIMILARITY.  
 FT ACT\_SITE 359 359 BY SIMILARITY.  
 FT ACT\_SITE 54 L -> I (IN REF. 2).  
 FT ACT\_SITE 64 I -> L (IN REF. 2).  
 FT ACT\_SITE 79 S -> D (IN REF. 2).  
 FT ACT\_SITE 84 G -> S (IN REF. 2).  
 FT ACT\_SITE 84 G -> S (IN REF. 2).  
 SQ SEQUENCE 514 AA; 58403 MW; A37712F3 CRC32;  
 Query Match 66.1%; Score 2559; DB 1; Length 514;  
 Best Local Similarity 64.0%; Pred. No. 0.00e+00;  
 Matches 330; Conservative 90; Mismatches 89; Indels 7; Gaps 5;

Db 1 MIQRKRTVFRVLNCTLLF--VSLPI-TKTSAYNGTLMQYFEWYTPNDQHKRLQND 57  
 QY 2 KLNRIISVLLTLLAVLFFPTEPAQAHHNGTNGTMQYFEWHLNDGNHNRRLRDD 61  
 Db 58 AEHLSDIGITAVWIPPAKGLSDNGYGPYDLYDLGEFOQKGTVRYTKYGTSELODAIG 117  
 QY 62 AANLKSIGITAVWIPPAWKTSGNDVYGAYDLYDLGEFQKGTVRYTKYGTSELODAIG 121  
 Db 118 SLHRSNVQYGVNHLNKHAGADATEDVAVENPANNRNOETSEYQIKAWTDREPRGRN 177  
 QY 122 SLKNGIOVYGVNMHKGADGTEMVNAVEVRNRSNRNQEISGEYITEAMTKDFPGRGH 181  
 Db 178 TYSDFKRWYHFDGADWDESKRI-SRIKPFEGEKANDWEVSSSENGNYDLYADYVDH 236  
 QY 182 THSNFKRWYHFDGADWDESKRI-SRIKPFEGEKANDWEVSSSENGNYDLYADYVDH 241  
 Db 237 PDVVAETKKGWIYANELSLDGFRIIDAAKHIFSLRDWQVAVQATGKEMFTVAEYWN 296  
 QY 242 PEVINELRNMGWYITNLNLDGFRIDAVKHISYTRDMLTHVRNWTGKPMFAEAFKWN 301  
 Db 297 NAGLENYLNKTSWNHVSFVDPVLYLYNASGGYFDMRNLNGSVVQKHPHATVFDV 356  
 QY 302 DLAAIENLNKTSWNHVSFVDPVLYLYNASGGYFDMRNLNGSVVQKHPHATVFDV 361  
 Db 357 NHDTPQGSLESTVQWFKFLAYAFILTRREGYPOVFGDMYGTGKTSPEISLKNONIE 416  
 QY 362 NHDSPQGSLESTVQWFKFLAYAFILTRREGYPOVFGDMYGTGKTSPEISLKNONIE 418  
 Db 417 PILKAREYAGPOHDIIDHPDVTGWTREGDSSAAKSLAALITDPGGSKRMVAGLNA 476  
 QY 419 PLLQARQTYAYGTQHDYFDHDDIIIGWTRGESSHPNSGLATINSDPGGNKMYVGHKKA 478  
 Db 477 GETWYDITGNRSDTYKIGSDGSGEHHVNDGVSIIYV 512  
 QY 479 GQWRDITGNRSGTIVNADGNGFTVNGGAVSVWV 514

RESULT 5  
 ID AMY2.SALTY STANDARD; PRT; 494 AA.  
 AC P26613;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE CYTOPLASMIC ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
 DE GLUCANOHYDROLASE).  
 GN AMYA.  
 OS SALMONELLA TYPHIMURIUM.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SJW1103;  
 RX MEDLINE; 93015717.  
 RA RAHA M., KAWAGISHI I., MUELLER V., KIHARA M., MACNAB R.M.;  
 RL J. BACTERIOL. 174:6644-6652(1992).  
 RN [2]  
 RP SEQUENCE OF 1-6 FROM N.A.  
 RC STRAIN-SJW1103;  
 RX MEDLINE; 92407478.  
 RA KAWAGISHI I., MUELLER V., WILLIAMS A.W., IRIKURA V.M., MACNAB R.M.;



[illegible]

|          |   |   |              |
|----------|---|---|--------------|
| Db       | 181   | YTGEGWQDQVDELGNFDYLMGENIDFRNHAVTEEIKYWARWVMEQTQCDGFLDQVAKHI | 240          |
| Qy       | 214   | -TGRADWEVDIENGNYDYILMYADIDMDHPVINELRNNGVYTNLNLDDGFRIDAVKHI  | 272          |
| Db       | 241   | PAWYKKEWIEHQVEAPKPLFVIAEYWSHEVDKLQYIIDOVEGKTMFLDAPLQMKFHEAS | 300          |
| Qy       | 273   | KYSTRWLTHVRNWTGKPMFAEAFKWNDLAAENYLNKTSWNHVSFDPVPLHYNLYNAS   | 332          |
| Db       | 301   | RMGRDYDWTQIFTGTLVEADPFHAVTLVANHDTPLQALEAPVPEWFKPLAYALILLREN | 360          |
| Qy       | 333   | NSGGYFDMRNLNGSVQKHPHIAHVFVDNHDSPQGEALSFVQSFKPLAYALILITREQ   | 392          |
| Db       | 361   | GVPSVFYDPIGAHYEDVGGGQTPIDMPIIEQLDELILARQFAHGVTOTLFFDHPNCI   | 420          |
| Qy       | 393   | GYPVSFVYGD-Y--Y-GIPTHG-V-P-SMK--SKIDPLQARQTYAYGTQHDYFDHDDII | 442          |
| Db       | 421   | AFSRSGTDEPGC-VV-VMSGDDGKTHLGENYGNKTRDFLGNRQERVYTDENGAT      | 478          |
| Qy       | 443   | GWTRREGDSHPNSGLATIMSDGPGGNKMYVGKHKAGQVWRDITGNRSGTVTINADGWN  | 502          |
| Db       | 479   | FFCNGGVSVMV 490   |              |
| Qy       | 503   | FTVNGGAVSMV 514   |              |
| RESULT 7 |   |   |              |
| ID       | AMTA_PSEST  | STANDARD;   | PRT; 548 AA. |
| AC       | PI3507;   |   |              |
| DT       | 01-JAN-1990 (REL. 13, CREATED)  |   |              |
| DT       | 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)                           |   |              |
| DT       | 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)                         |   |              |
| DE       | GLUCAN 1,4-ALPHA-MALTOTETRAHYDROLASE PRECURSOR (EC 3.2.1.60) (G4-     |   |              |
| DE       | AMYLASE) (MALTOTETRAOSE-FORMING AMYLASE) (EXO-MALTOTETRAOCHYDROLASE)  |   |              |
| DE       | (MALTOTETRAOSE-FORMING EXO-AMYLASE).                                  |   |              |
| GN       | AMYP.   |   |              |
| OS       | PSUDOMONAS STUTZERI (PSUDOMONAS PERFECTOMARINA).                      |   |              |
| OS       | PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;      |   |              |
| OC       | PSUDOMONADACEAE.  |   |              |
| RC       | [1]   |   |              |
| RC       | SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.                             |   |              |
| RP       | STRAIN-MO-19;   |   |              |
| RX       | MEDLINE; 89155431.  |   |              |
| RA       | FUJITA M., TORIGOE K., NAKADA T., TSUSAKI K., KUBOTA M., SAKAI S.,    |   |              |
| RA       | TSUJISAKA Y.;   |   |              |
| RL       | J. BACTERIOL. 171:1333-1339(1989).                                    |   |              |
| RL       | [2]   |   |              |
| RP       | X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND REVISIONS TO 286-302.      |   |              |
| RP       | STRAIN-MO-19;   |   |              |
| RX       | MEDLINE; 97271999.  |   |              |
| RA       | MORISHITA Y., HASEGAWA K., MATSUURA Y., KATSUBE Y., KUBOTA M.,        |   |              |
| RA       | SAKAI S.;   |   |              |
| RL       | J. MOL. BIOL. 267:661-672(1997).                                      |   |              |
| RL       | [3]   |   |              |
| RP       | X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF VARIANT GLN-240.             |   |              |
| RP       | STRAIN-MO-19;   |   |              |
| RX       | MEDLINE; 97428332.  |   |              |
| RA       | YOSHIOKA Y., HASEGAWA K., MATSUURA Y., KATSUBE Y., KUBOTA M.;         |   |              |
| RL       | J. MOL. BIOL. 271:619-628(1997).                                      |   |              |
| CC       | -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-ALPHA-D-GLUCOSIDIC LINKAGES |   |              |
| CC       | IN AMYLACEOUS POLYSACCHARIDES SO AS TO REMOVE SUCCESSIVE              |   |              |
| CC       | MALTOTETRAOSE RESIDUES FROM THE NON-REDUCING CHAIN ENDS.              |   |              |
| CC       | -!- COFACTOR: BINDS TWO CALCIUM IONS.                                 |   |              |
| CC       | -!- PATHWAY: DEGRADATION OF STARCH.                                   |   |              |
| CC       | -!- SUBCELLULAR LOCATION: EXTRACELLULAR.                              |   |              |
| CC       | -!- THERE ARE SEVERAL ISOENZYME FORMS OF THIS PROTEIN.                |   |              |
| CC       | -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO     |   |              |
| CC       | KNOWN AS THE ALPHA-AMYLASE FAMILY.                                    |   |              |
| CC       | EMBL; M24516; G151013; -  |   |              |
| DR       | PUR; A32803; A32803   |   |              |
| DR       | PDB; 2AMG; 01-APR-97.   |   |              |
| DR       | PDB; 1JDA; 15-OCT-97.   |   |              |
| DR       | PDB; 1JDC; 15-OCT-97.   |   |              |
| DR       | PDB; 1JDD; 15-OCT-97.   |   |              |



```

Db 330 YGLDMSLSTADYKSVNDQVTFDNDHMDRFOVSGANGRK-LEQ-ALALTLTS-RCVPA 386
Qy 338 FDMRNLNGSVVQKHPIHA-VTFVDNHDSPGGEALSFVQSWFKPLALAILITREQYPS 396
Db 387 IYIGTEQY-MTNGGDPNNRAKMSFSFSTSTAY 417
Qy 397 VFYIG-DYIGIPTHGVPSMKSIDPLQARQTY 427

RESULT 10
ID AMYR_BACS8 STANDARD; PRT; 713 AA.
AC P17692;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE RAW-STARCH-DIGESTING AMYLASE PRECURSOR (EC 3.2.1.1) (1.4-ALPHA-D-
DE GLUCAN GLUCANOHYDROLASE).
OS BACILLUS SP. (STRAIN B1018).
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-47.
RX MEDLINE; 90147765.
RL ITOR P., TSUKAGOSHI N., UDAKA S.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 166:630-636(1990).
CC -!- FUNCTION: THIS ENDO-TYPE ADSORBABLE AMYLASE IS CAPABLE TO
CC DIGEST RAW-STARCH.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -!- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR EMBL; M33302; G142509; -.
DR EMBL; D90112; G216325; -.
DR PIR; A34648; A34648.
DR PIR; S09196; S09196.
DR HSP; P43379; ICDG.
KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; CALCIUM; SIGNAL.
FT SIGNAL 1 27
FT CHAIN 28 713 RAW-STARCH DIGESTING AMYLASE.
FT ACT SITE 256 256 BY SIMILARITY.
FT ACT SITE 260 260 BY SIMILARITY.
FT ACT SITE 355 355 BY SIMILARITY.
FT ACT SITE 355 355 BY SIMILARITY.
SQ SEQUENCE 713 AA; 77420 MW; 5C83248D CRC32;

Query Match 5.0%; Score 192; DB 1; Length 713;
Best Local Similarity 25.5%; Pred. No. 1.30e-15;
Matches 54; Conservative 64; Mismatches 75; Indels 19; Gaps 16;

Db 214 ENGIYKNL-YDLADLNHNNSVDVYLKDAIKMWL-D-LGIDGIRMDAVKHPFGWQKSF 270
Qy 225 ENGNIDYLMYADIDMDHPEVINE--LRNW-GVWYTNLTLDGFRIDAVKHKISYTRDWL 281
Db 271 AAV-NNY-KPVTFGEWFLGVNEVGENHKFANESGMSLLDFFPAQKRVFRD-NYDNM 327
Qy 282 THVRNTTGKPMFAVAF-FWK-NDLAAIEN-YLNKTSWNHVSFVDVPLHY-NLYNASNGGY 337
Db 328 YGLKAMLEGAADYAOVDQVTFIDNHDMEHFASNANRKLKLEALFTLILARV---PA 384
Qy 338 FDMRNLNGSVVQKHPIHA-VTFVDNHDSPGGEALSFVQSWFKPLALAILITREQYPS 396
Db 385 IYIGTEQY-MSGGTDPDNRARIPFSFSTSTAY 415
Qy 397 VFYIG-DYIGIPTHGVPSMKSIDPLQARQTY 427

RESULT 11
ID CDGT_BACS3 STANDARD; PRT; 712 AA.
AC P09121;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE CYCLOMALTODEXTRIN GLUCANOTRANSFERASE PRECURSOR (EC 2.4.1.19)

```

```

DE (CYCLODEXTRIN-GLYCOSYLTRANSFERASE) (CGTASE).
GN BACILLUS SP. (STRAIN 38-2).
OS PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-44.
RX MEDLINE; 89036108.
RA KANEKO T., HAMAMOTO T., HORIKOSHI K.;
RL J. GEN. MICROBIOL. 134:97-105(1988).
RN [2]
RP SEQUENCE OF 1-586 FROM N.A.
RA HAMAMOTO T., KANEKO T., HORIKOSHI K.;
RL AGRIC. BIOL. CHEM. 51:2019-2022(1987).
CC -!- CATALYTIC ACTIVITY: DEGRADE STARCH TO CYCLODEXTRINS BY FORMATION
CC OF A 1,4-ALPHA-D-GLUCOSIDIC BOND.
CC -!- COFACTOR: BINDS TWO CALCIUM IONS.
CC -!- SUBUNIT: MONOMER.
CC -!- CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE IN THE AMINO-
CC TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND IN STARCH,
CC AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER ACTIVITIES,
CC INCLUDING THE RECONSTITUTION OF AN ALPHA-1,4-GLUCOSIDIC LINKAGE
CC FOR CYCLIZING THE MALTOOLIGOSACCHARIDE PRODUCED.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR EMBL; M19880; G142678; -.
DR EMBL; D00129; G216248; -.
DR PIR; S24193; ALBSG3.
DR HSP; P43379; ICDG.
KW TRANSFERASE; GLYCOSYLTRANSFERASE; CALCIUM; SIGNAL.
FT SIGNAL 1 27
FT CHAIN 28 712 CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
FT DOMAIN 28 165 A1.
FT DOMAIN 166 229 B.
FT DOMAIN 230 433 A2.
FT DOMAIN 434 522 C.
FT DOMAIN 523 608 D.
FT DOMAIN 609 712 E.
FT DISULFID 70 77 BY SIMILARITY.
FT ACT SITE 256 256 BY SIMILARITY.
FT ACT SITE 284 284 BY SIMILARITY.
FT ACT SITE 355 355 BY SIMILARITY.
FT ACT SITE 355 355 BY SIMILARITY.
FT CONFLICT 582 586 VPGGI -> SWRHL (IN REF. 2).
SQ SEQUENCE 712 AA; 78249 MW; 6B5DE548 CRC32;

Query Match 4.9%; Score 189; DB 1; Length 712;
Best Local Similarity 25.7%; Pred. No. 4.62e-15;
Matches 55; Conservative 61; Mismatches 77; Indels 21; Gaps 20;

Db 213 IENGIYKNL-YDLADLNHNNSVDVYLKDAIKMWL-D-LGVGIRVDVAKHMPFGWQKSF 269
Qy 224 IENGNIDYLMYADIDMDHPEV-INE-LRNW-GVWYTNLTLDGFRIDAVKHKISYTRDW 280
Db 270 MSTI-NNY-KPVTFGEWFLGVNEISPHYHOFANESGMSLLDFFPAQKARQVFRD-NTD 325
Qy 281 LTHVRNTTGKPMFAVAFEFWKNDLAAIE-NYLN-KTSWNHVSFVDVPL-H-Y-NLYNASNG 335
Db 326 NMYGLKAMLEGEVDYAOVNDQVTFIDNHDMEHFHNSGDRRLKLEALFTLA-SR--GV 382
Qy 336 GYFDMRNLNGSVVQKHPIHA-VTFVDNHDSPGGEALSFVQSWFKPLALAILITREQY 394
Db 383 PAIYIGSEQY-MSGGNDPDNRARIPFSFSTTTAY 415
Qy 395 PSVFYIG-DYIGIPTHGVPSMKSIDPLQARQTY 427

RESULT 12
ID CDGT_BACS3 STANDARD; PRT; 718 AA.
AC P31747;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE CYCLOMALTODEXTRIN GLUCANOTRANSFERASE PRECURSOR (EC 2.4.1.19)
DE (CYCLODEXTRIN-GLYCOSYLTRANSFERASE) (CGTASE).

```

GN CGT.  
OS BACILLUS SP. (STRAIN 6.6.3).  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA AKHMETZJANOV A.A.;  
RL SUBMITTED (MAY-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- CATALYTIC ACTIVITY: DEGRADE STARCH TO CYCLODEXTRINS BY FORMATION  
CC OF A 1,4-ALPHA-D-GLUCOSIDIC BOND.  
CC -!- COFACTOR: BINDS TWO CALCIUM IONS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE IN THE AMINO-  
CC TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND IN STARCH,  
CC AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER ACTIVITIES,  
CC INCLUDING THE RECONSTITUTION OF AN ALPHA-1,4-GLUCOSIDIC LINKAGE  
CC FOR CYCLIZING THE MALTOOLIGOSACCHARIDE PRODUCED.  
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
DR EMBL; X66106; G39839; -.  
DR PIR; S21532; ALBSG6.  
DR HSSP; P30920; ICGT.  
KW TRANSFERASE; GLYCOSYLTRANSFERASE; CALCIUM; SIGNAL.  
FT SIGNAL 1 34  
FT CHAIN 35 718  
FT DOMAIN 35 172  
FT DOMAIN 173 236  
FT DOMAIN 237 440  
FT DOMAIN 441 528  
FT DOMAIN 529 614  
FT DOMAIN 615 718  
FT DISULFID 77 84  
FT ACT\_SITE 263 263  
FT ACT\_SITE 291 291  
FT ACT\_SITE 362 362  
SQ SEQUENCE 718 AA; 73784DA0 CRC32;  
  
Query Match 4.9%; Score 189; DB 1; Length 718;  
Best Local Similarity 24.2%; Pred. No. 4.62e-15;  
Matches 52; Conservative 69; Mismatches 71; Indels 23; Gaps 21;  
  
Db 220 LENGIYKNL-YDLADFNHNNTIDKYFKAIKLWL-D-MGVGIRVDAVKHIALGWQSW 276  
QY 224 IENGNDYLMYADIMDHPEV-INE-LRNW-GWYNTNLNLDGFRIDAVKHKYSYTRDW 276  
Db 277 MSSI-IVH-KPVFTGEGVGLSAAADNTDFANKSGMSLLDFRNSAVR-NVFRDNTSN 333  
QY 281 LTHVRNTTKPMFAVAE-FWKNDLAIEN--YLNKTSWNH-SV-FDVPPLHYNLNYSNSG 335  
Db 334 MY-ALDSMINSTATYNOVNDQVTFIDNHDMDRFT-SAVNNRLEQALAFILT-SR--G 388  
QY 336 GYFDMRNILGVSQKPIHA-VTFVDNHDSPQGEALSFQSWF-KPLAYAILITREQ 393  
Db 389 VPAIYIGTEQY-LTGNGDDPNRAKHPFSKSTTAF 422  
QY 394 YPSVFYG-DYIGTHGVFSKSKIDPLQARQTY 427  
  
RESULT 13  
ID CCGU\_BACCI STANDARD; PRT; 713 AA.  
AC P43379;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE CYCLOMALTOEXTRIN GLUCANOTRANSFERASE (UPDATE)  
DE (CYCLODEXTRIN-GLYCOSYLTRANSFERASE, PRECURSOR (EC 2.4.1.19))  
OS BACILLUS CIRCULANS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RC STRAIN-251;  
RX MEDLINE; 94149761.  
RA LAWSON C.L., DE VRIES G.E., PENNINGA D., DIJKHUIZEN L., ROZEBOOM H.J.,  
RA KALK K.H., KALK K.H., STROKOPYTOV B., DIJKHUIZEN L., DIJKSTRA B.W.,

J. MOL. BIOL. 236:590-600(1994).  
RL [2]  
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
RC STRAIN-251;  
RX MEDLINE; 96094317.  
RA KNEGTEL R.M.A., STROKOPYTOV B., PENNINGA D., FABER O.G.,  
RA ROZEBOOM H.J., KALK K.H., DIJKSTRA B.W., DIJKHUIZEN L.,  
RL J. BIOL. CHEM. 270:29256-29264(1995).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE; 97115811.  
RA PENNINGA D., VAN DER VEEN B.A., KNEGTEL R.M., VAN HIJUM S.A.,  
RA ROZEBOOM H.J., KALK K.H., DIJKSTRA B.W., DIJKHUIZEN L.,  
RL J. BIOL. CHEM. 271:32777-32784(1996).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH INHIBITOR.  
RC STRAIN-251;  
RX MEDLINE; 96264806.  
RA STROKOPYTOV B., KNEGTEL R.M.A., PENNINGA D., ROZEBOOM H.J., KALK K.H.,  
RA DIJKHUIZEN L., DIJKSTRA B.W.;  
RL BIOCHEMISTRY 35:4241-4249(1996).  
CC -!- CATALYTIC ACTIVITY: DEGRADE STARCH TO CYCLODEXTRINS BY FORMATION  
CC OF A 1,4-ALPHA-D-GLUCOSIDIC BOND.  
CC -!- COFACTOR: BINDS TWO CALCIUM IONS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- DOMAIN: MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE IN THE AMINO-  
CC TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND IN STARCH, AND  
CC THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER ACTIVITIES,  
CC INCLUDING THE RECONSTITUTION OF AN ALPHA-1,4-GLUCOSIDIC LINKAGE  
CC FOR CYCLIZING THE MALTOOLIGOSACCHARIDE PRODUCED.  
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
DR EMBL; X78145; G510492; -.  
DR PDB; 1CXG; 20-DEC-94.  
DR PDB; 1CDG; 08-MAR-95.  
DR PDB; 1CGV; 27-FEB-95.  
DR PDB; 1CGW; 27-FEB-95.  
DR PDB; 1CGX; 07-FEB-95.  
DR PDB; 1CGY; 07-FEB-95.  
DR PDB; 1CXE; 15-DEC-95.  
DR PDB; 1CAF; 15-DEC-95.  
DR PDB; 1CXH; 15-DEC-95.  
DR PDB; 1CXI; 15-DEC-95.  
DR PDB; 1DIJ; 21-APR-97.  
DR PDB; 1TCM; 21-APR-97.  
KW TRANSFERASE; GLYCOSYLTRANSFERASE; CALCIUM; SIGNAL; 3D-STRUCTURE.  
FT SIGNAL 1 25  
FT CHAIN 26 713  
FT DOMAIN 28 165  
FT DOMAIN 166 229  
FT DOMAIN 230 433  
FT DOMAIN 434 522  
FT DOMAIN 523 609  
FT DOMAIN 610 713  
FT DISULFID 70 77  
FT ACT\_SITE 256 256  
FT ACT\_SITE 284 284  
FT ACT\_SITE 355 355  
SQ SEQUENCE 713 AA; 77309 MW; 3DC02138 CRC32;  
  
Query Match 4.8%; Score 185; DB 1; Length 713;  
Best Local Similarity 25.0%; Pred. No. 2.48e-14;  
Matches 53; Conservative 64; Mismatches 76; Indels 19; Gaps 18;  
  
Db 214 ENGIYKNL-YDLADFNHNNTVDYVYKDAIKWL-D-LGIDGIRMDAVKHPFGQKSF 270  
QY 225 ENGNIDYLMYADIMDHPE-VINE-LRNW-GWYNTNLNLDGFRIDAVKHKYSYTRDW 270  
Db 271 AAV-NNY-KPVFTGEGVGLGVSNEPNSHNFANESGMSLLDFRFAQKRVQFRD-NTDNM 281  
QY 282 THVRNTTKPMFAVAE-FWKNDLAIEN-YLNKTSWNHVSFVDVPLHY-NLYNNSGGY 337  
Db 338 YGLKAMLEGSADYQVDDQVTFIDNHDMDRFT-SAVNNRLEQALAFILT-SR--GVPA 384

|    |        |     |     |
|----|--------|-----|-----|
| FT | HELIX  | 88  | 98  |
| FT | TURN   | 97  | 98  |
| FT | HELIX  | 99  | 103 |
| FT | TURN   | 104 | 104 |
| FT | STRAND | 107 | 109 |
| FT | STRAND | 114 | 116 |
| FT | STRAND | 121 | 123 |
| FT | TURN   | 124 | 125 |
| FT | STRAND | 126 | 128 |
| FT | TURN   | 131 | 132 |
| FT | STRAND | 136 | 142 |
| FT | TURN   | 144 | 146 |
| FT | HELIX  | 149 | 161 |
| FT | TURN   | 162 | 163 |
| FT | STRAND | 165 | 170 |
| FT | TURN   | 172 | 173 |
| FT | STRAND | 174 | 177 |
| FT | TURN   | 183 | 184 |
| FT | TURN   | 186 | 189 |
| FT | STRAND | 191 | 193 |
| FT | TURN   | 194 | 195 |
| FT | STRAND | 196 | 199 |
| FT | TURN   | 202 | 203 |
| FT | TURN   | 205 | 206 |
| FT | STRAND | 209 | 210 |
| FT | STRAND | 214 | 215 |
| FT | HELIX  | 220 | 224 |
| FT | STRAND | 227 | 228 |
| FT | TURN   | 229 | 230 |
| FT | STRAND | 231 | 234 |
| FT | TURN   | 236 | 237 |
| FT | HELIX  | 239 | 243 |
| FT | TURN   | 255 | 256 |
| FT | STRAND | 259 | 262 |
| FT | HELIX  | 265 | 267 |
| FT | HELIX  | 270 | 281 |
| FT | TURN   | 282 | 283 |
| FT | STRAND | 287 | 290 |
| FT | HELIX  | 301 | 309 |
| FT | STRAND | 313 | 315 |
| FT | HELIX  | 317 | 327 |
| FT | TURN   | 328 | 329 |
| FT | HELIX  | 334 | 347 |
| FT | TURN   | 339 | 349 |
| FT | HELIX  | 351 | 353 |
| FT | STRAND | 355 | 356 |
| FT | TURN   | 361 | 362 |
| FT | TURN   | 369 | 370 |
| FT | HELIX  | 373 | 385 |
| FT | STRAND | 389 | 393 |
| FT | TURN   | 396 | 399 |
| FT | TURN   | 405 | 408 |
| FT | HELIX  | 420 | 428 |
| FT | TURN   | 429 | 430 |
| FT | HELIX  | 431 | 434 |
| FT | HELIX  | 436 | 440 |
| FT | STRAND | 442 | 449 |
| FT | TURN   | 450 | 451 |
| FT | STRAND | 452 | 459 |
| FT | TURN   | 460 | 461 |
| FT | STRAND | 462 | 469 |
| FT | STRAND | 476 | 478 |
| FT | STRAND | 482 | 482 |
| FT | STRAND | 488 | 490 |
| FT | TURN   | 493 | 498 |
| FT | STRAND | 503 | 506 |
| FT | TURN   | 507 | 508 |
| FT | STRAND | 509 | 510 |
| FT | STRAND | 513 | 515 |
| FT | TURN   | 517 | 518 |
| FT | STRAND | 520 | 525 |
| FT | STRAND | 533 | 538 |
| FT | STRAND | 541 | 542 |

US-08-952-741-2.rsp

Wed Sep 16 09:49:09 1998

```

CC      544      545
CC      547      553
CC      561      564
CC      565      566
CC      569      569
CC      571      572
CC      574      578
CC      579      580
CC      581      585
CC      592      600
CC      601      602
CC      603      613
CC      618      626
CC      632      633
CC      635      640
CC      643      645
CC      646      646
CC      652      653
CC      655      655
CC      659      659
CC      666      667
CC      677      678
CC      680      688
CC      693      695
CC      701      704
CC      710      716
CC      718 AA; 78046 MW; CF4C4096 CRC32;
SQ      SEQUENCE 718 AA; 78046 MW; 4.8%; Score 186; DB 1; Length 718;
Query Match 4.8%; Score 186; DB 1; Length 718;
Best Local Similarity 23.7%; Pred. No. 1.63e-14; Indels 23; Gaps 21;
Matches 51; Conservative 71; Mismatches 70;
Db      220 LENGIYKNI-YDLADFNINNFATIDKYFKDAIKLWL-D-MGVGIRVDVAKHMPGLGWSW 276
QY      224 IENGNYDILMYADIDMDHPEV-INE-LRNW-GVWYTNLNLGDFRIDAVKHKIKSYTRDW 280
Db      277 MSSI-YAH-KPVFTFGFWELGSAADADNTDFANKSGMSLLDFRFSNVR-NVFRONTSN 333
QY      281 LTHVRNTGKPMFAVAE-EWKNDLAAIEN--YLNKTSNH-SV-FDVPPLHYNLYNASNG 335
Db      334 MY-ALDSMINSTATNOVNDQVTFIDNHDMDRFT-SAVNNRLEQALFTLT-SR-G 388
QY      336 GFYDMRNLNGSVVQKPIHA-VTFVDNHDSPGSALESFVQSWE-KPLAVALILTREQ 393
Db      389 VPALYGTGEQY-LTGNGDPPNRKMPSEKSTTAF 422
QY      394 YPSVFG-DYVGIPTHGVPSMKSKIDPLLOARQTY 427

RESULT 15 STANDARD; PRT; 711 AA.
ID      CDGT_BACST
AC      P31797;
DT      01-JUL-1993 (REL. 26, CREATED)
DT      01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT      01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DT      01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE      CYCLOMALTODEXTRIN GLUCANOTRANSFERASE (CGTASE).
DE      (CYCLODEXTRIN-GLYCOSYLTRANSFERASE)
GN      BACILLUS STEAROTHERMOPHILUS.
OS      PRORARIOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
OC      [1]
RN      SEQUENCE FROM N.A.
RP      STRAIN-NO. 2;
RC      MEDLINE; 93119155.
RA      FUJIWARA S., KANEMOTO M., KIM B., LEJEUNE A., SAKAGUCHI K.,
RA      IMANAKA T.;
RA      APPL. ENVIRON. MICROBIOL. 58:4016-4025(1992).
RL      [2]
RN      SEQUENCE FROM N.A., AND SEQUENCE OF 32-41.
RP      SUGIMOTO T., KUBOTA M., SAKAI S.;
RL      PATENT NUMBER UK2169902.
CC      -!- CATALYTIC ACTIVITY: DEGRADE STARCH TO CYCLODEXTRINS BY FORMATION

```

Search completed: Tue Sep 15 14:19:18 1998  
Job time : 63 secs.

\*\*\*\*\*  
WILEY  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 15 14:19:35 1998; MasPar time 31.41 Seconds  
691.851 Million cell updates/sec

Tabular output not generated.

Title: >US-08-952-741-2  
Description: (1-516) from US08952741.pep  
Perfect Score: 3873  
Sequence: 1 MKLHNRIISVLLTLLAVAV.....ADGWGFTVNGGAVSVYVVKQ 516

Scoring table: PAM 150  
Gap 11

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sptrembl6  
1:sp\_fungi 2:sp\_human 3:sp\_invertebrate 4:sp\_mammal  
5:sp\_nhc 6:sp\_organelle 7:sp\_phase 8:sp\_plant  
9:sp\_bacteria 10:sp\_rodent 11:sp\_virus 12:sp\_vertebrate  
13:sp\_unclassified

Statistics: Mean 49.603; Variance 91.345; scale 0.543

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES  |       |                       |  |
|------------|-------|-----------------------|--|
| Result No. | Score | Query Match Length DB | ID Description Pred. No.                     |
| 1          | 2672  | 69.0                  | 613 9 Q59222 ALPHA-AMYLASE (EC 3.2. 0.00e+00 |
| 2          | 2648  | 68.4                  | 521 9 P71034 ALPHA-AMYLASE PRECURSO 0.00e+00 |
| 3          | 2642  | 68.2                  | 549 9 Q31193 ALPHA-AMYLASE. 0.00e+00         |
| 4          | 2608  | 67.3                  | 549 9 Q45519 ALPHA AMYLASE PRECURSO 0.00e+00 |
| 5          | 1912  | 49.4                  | 423 9 Q03657 AMYLASE (EC 3.2.1.-). 0.00e+00  |
| 6          | 1682  | 43.4                  | 485 9 Q53786 AMYLASE. 0.00e+00               |
| 7          | 388   | 9.5                   | 461 9 Q33476 ALPHA-AMYLASE PRECURSO 3.77e-49 |
| 8          | 362   | 9.3                   | 460 9 Q08452 ALPHA-AMYLASE (EC 3.2. 6.41e-48 |
| 9          | 172   | 4.4                   | 504 9 Q60224 ALPHA-AMYLASE PRECURSO 2.05e-11 |
| 10         | 150   | 3.9                   | 482 9 Q60051 ALPHA-AMYLASE PRECURSO 9.61e-08 |
| 11         | 148   | 3.8                   | 423 8 Q42678 ALPHA-AMYLASE PRECURSO 2.02e-07 |
| 12         | 135   | 3.5                   | 407 8 Q41442 ALPHA-AMYLASE PRECURSO 2.27e-05 |
| 13         | 135   | 3.5                   | 427 8 Q03651 ALPHA-AMYLASE TYPE B P 2.27e-05 |
| 14         | 130   | 3.4                   | 349 8 Q41441 ALPHA-AMYLASE (FRAGMEN 1.31e-04 |
| 15         | 132   | 3.4                   | 421 8 Q42504 ALPHA-AMYLASE TYPE A. 6.53e-05  |
| 16         | 131   | 3.4                   | 427 8 Q40015 ALPHA-AMYLASE 1. 9.26e-05       |
| 17         | 130   | 3.4                   | 437 8 Q04965 ALPHA-AMYLASE (EC 3.2. 1.31e-04 |
| 18         | 131   | 3.4                   | 438 8 Q40017 ALPHA-AMYLASE 2. 9.26e-05       |
| 19         | 128   | 3.3                   | 429 8 Q40016 ALPHA-AMYLASE 1. 2.62e-04       |
| 20         | 127   | 3.3                   | 499 1 Q00250 TAKA-AMYLASE A (TAA-G1 3.70e-04 |

|    |     |     |      |    |        |                                 |          |
|----|-----|-----|------|----|--------|---------------------------------|----------|
| 21 | 129 | 3.3 | 1590 | 9  | Q59983 | GLUCOSYLTRANSFERASE-I           | 1.86e-04 |
| 22 | 129 | 3.3 | 1590 | 9  | Q55263 | GLUCOSYLTRANSFERASE GT          | 1.86e-04 |
| 23 | 125 | 3.2 | 437  | 8  | Q04964 | ALPHA-AMYLASE (EC 3.2. 7.33e-04 |          |
| 24 | 122 | 3.2 | 640  | 1  | Q13296 | ACID-STABLE ALPHA-AMYL          | 2.02e-03 |
| 25 | 120 | 3.1 | 439  | 8  | Q41770 | ALPHA-AMYLASE                   | 3.93e-03 |
| 26 | 121 | 3.1 | 513  | 1  | Q14154 | ALPHA AMYLASE PRECURSO          | 2.82e-03 |
| 27 | 120 | 3.1 | 552  | 9  | Q44528 | PUTATIVE ALPHA-GLUCANO          | 3.93e-03 |
| 28 | 119 | 3.1 | 700  | 9  | Q24781 | ALPHA-AMYLASE PRECURSO          | 5.48e-03 |
| 29 | 116 | 3.0 | 441  | 9  | Q60035 | 4-ALPHA-GLUCANOTRANSFE          | 1.46e-02 |
| 30 | 116 | 3.0 | 475  | 9  | Q56791 | ALPHA-AMYLASE PRECURSO          | 1.46e-02 |
| 31 | 115 | 3.0 | 570  | 1  | Q01117 | ALPHA-AMYLASE.                  | 2.02e-02 |
| 32 | 115 | 3.0 | 590  | 9  | Q45490 | BSMA.                           | 2.02e-02 |
| 33 | 118 | 3.0 | 725  | 9  | Q59239 | CYCLODEXTRIN GLUCANOTR          | 7.62e-03 |
| 34 | 114 | 2.9 | 437  | 8  | Q40018 | ALPHA-AMYLASE 2.                | 2.79e-02 |
| 35 | 111 | 2.9 | 466  | 9  | Q52414 | ALPHA-AMYLASE PRECURSO          | 7.26e-02 |
| 36 | 113 | 2.9 | 614  | 9  | Q52516 | MALTOPEPTIDASE FORMING          | 3.85e-02 |
| 37 | 113 | 2.9 | 652  | 9  | Q59242 | 1,4-ALPHA-GLUCAN BRANC          | 3.85e-02 |
| 38 | 113 | 2.9 | 1290 | 9  | Q48756 | DEXTRANSUCRASE.                 | 3.85e-02 |
| 39 | 114 | 2.9 | 1449 | 9  | Q55264 | GLUCOSYLTRANSFERASE PR          | 2.79e-02 |
| 40 | 107 | 2.8 | 381  | 9  | Q52026 | SIGNAL-TRANSDUCING PRO          | 2.53e-01 |
| 41 | 108 | 2.8 | 421  | 9  | Q59243 | TRUNCATED 1,4-ALPHA-GL          | 1.86e-01 |
| 42 | 108 | 2.8 | 491  | 1  | Q13996 | AMYLASE PRECURSOR.              | 1.86e-01 |
| 43 | 107 | 2.8 | 507  | 1  | Q08806 | ALPHA-AMYLASE 2 (EC 3.          | 2.53e-01 |
| 44 | 107 | 2.8 | 1599 | 9  | Q00599 | GLUCOSYLTRANSFERASE S           | 2.53e-01 |
| 45 | 105 | 2.7 | 394  | 11 | Q04537 | GENOME, PARTIAL SEQUEN          | 4.65e-01 |

ALIGNMENTS

RESULT 1  
ID Q59222 PRELIMINARY; PRT; 613 AA.  
AC Q59222;  
DT 01-NOV-1996 (TREMUREL. 01, CREATED)  
DT 01-NOV-1996 (TREMUREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMUREL. 01, LAST ANNOTATION UPDATE)  
DE ALPHA-AMYLASE (EC 3.2.2.1) (PURINE NUCLEOSIDASE).  
GN AMY.  
OS BACILLUS SP.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TS-23;  
RA LIN L.L., CHU W.S., HSU W.H.;  
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- CATALYTIC ACTIVITY: A N-D-RIBOSYLPURINE + H(2)O -  
CC A PURINE + D-RIBOSE.  
DR EMBL; U22045; G722279; -.  
KW HYDROLASE; GLYCOSIDASE.  
SQ SEQUENCE 613 AA; 69537 MW; DC66587C CRC32;

|                       |     |  |     |  |
|-----------------------|-----|--|-----|--|
| Query Match           |     |  |     | 69.0%; Score 2672; DB 9; Length 613;                   |
| Best Local Similarity |     |  |     | 68.5%; Pred. No. 0.00e+00;                             |
| Matches               |     |  |     | 342; Conservative 78; Mismatches 76; Indels 3; Gaps 2; |
| Db                    | 17  | LVISFFTPFST--AQANTAPINETMMQYFEWDLNDGTLTKVKNAAANLSLGTALWL   | 74  |  |
| Qy                    | 16  | LAVAVLFPPYTPAQAHNGINGIMQYFEWHLNDGNHNRDLRDDAANLKSIGTAWI     | 75  |  |
| Db                    | 75  | PPAYKGTQSQDVGVYDLDGFEFNQKGTIRTKYGTQYQIAQAAKAGMQYADV        | 134 |  |
| Qy                    | 76  | PPAWKGTQNDVGYGAYDLDGFEFNQKGTIRTKYGTQYQIAQAAKAGMQYADV       | 135 |  |
| Db                    | 135 | FNHKGAGDGTFEVDVDPNRSNQTSGTQIQIANTKDFPGRGNTYSSPKRWYHFDG     | 194 |  |
| Qy                    | 136 | MNHKGAGDGTFEVDVDPNRSNQTSGTQIQIANTKDFPGRGNTYSSPKRWYHFDG     | 195 |  |
| Db                    | 195 | TDWDSRKL-NRIYKFRSTGKAWDEVTENGNDYLMFADLMDHPVETLKNWGTW       | 253 |  |
| Qy                    | 196 | TDWDSRQLNKIKYKFRGTGKAWDEVTENGNDYLMFADLMDHPVETLKNWGTW       | 255 |  |
| Db                    | 254 | VNTNIDGFRDVKHIIKYSFFPDMLTYYVRNQTGKNLFAVGFEFWSYVKNLHNYTKNG  | 313 |  |
| Qy                    | 256 | TNTLNLDGFRIDAVKHIIKYSYTRDMLTHVRNTTGKPMFAVEFKNLDAAIENLNKTSW | 315 |  |



Db 314 SMSLFDAPLHNNYTTASKSGGYEDMYRLNNTLMKQPSLAVTLVDNHDTPQGSLSQSV 373  
QY 316 NHSVDFVPLHYNLASNGGYEDMRNLNGSVVQKPHIAVTFVDNHDSPQCEALESEFV 375  
Db 374 EPWFKPLAYAFILTROEGYPCVFGYDYGTPKYNIPGLKSKIDPLLIARDYAYGTQDY 433  
QY 376 QSMFKPLAYALILTROEGYPSVFGYDYGTHGVFSMKSKIDPLLOARQTYAGTQDY 435  
Db 434 IDHQDIIGTWREGIDTKPNPSGLAALITDGPGGSKMYGKHKHAGKYFYDLTGNRSDTVTI 493  
QY 436 FDHDIIGTWREGDSSHPSGLATIMSDGPGKNMYGKHKAGQVWRDITGNRSCTVTI 495  
Db 494 NADGWGEFKYNGSVSIWV 512  
QY 496 NADGWGFTVNGGAVSVWV 514  
RESULT 2 PRELIMINARY; PRT; 521 AA.  
ID P71034;  
AC P71034;  
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
DE ALPHA-AMYLASE PRECURSOR.  
OS BACILLUS SP.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MK 716;  
RA SIDHU G.S., CHAKRABARTI T.;  
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: U79445; G1667474; -.  
KW SIGNAL.  
FT CHAIN 35 521  
SQ SEQUENCE 521 AA; 59311 MW; 55960B19 CRC32;  
Query Match 68.4%; Score 2648; DB 9; Length 521;  
Best Local Similarity 67.1%; Pred. No. 0.00e+00;  
Matches 339; Conservative 80; Mismatches 81; Indels 5; Gaps 5;  
Db 15 LLAFLITAS-LFCPTGQPAKAA-AFENGTMQYFEWLPDDGTLTWKVANEANLSSIGI 72  
QY 11 LTLTLLAVAVLPFYMTPEQAHHNGTNGTMQYFEWHLPDNGNHNRLRDAANLKSIGI 70  
Db 73 TALMLPPAYKTSRSDVGYDYLDLGEFNQKAVRTKYGTKAQYLO-AIOAAHAGMQ 131  
QY 71 TAVWIIPPAWKGTSONDVGYGAYLDLGEFNQKAVRTKYGTGRSQ-LQGAVTSLKNGIQ 129  
Db 132 VYADVVDHKGAGDTEWDAVEVNPDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 191  
QY 71 TAVWIIPPAWKGTSONDVGYGAYLDLGEFNQKAVRTKYGTGRSQ-LQGAVTSLKNGIQ 129  
Db 132 VYADVVDHKGAGDTEWDAVEVNPDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 191  
QY 130 VYGDVVMNHKGAGDTEWDAVEVNPDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 189  
Db 192 WYHFGDWDDESKL-SRIYKFRGIGKAWDEVDTENGNYDLYMADLMDHPEVVTTELK 250  
QY 190 WYHFGDWDQSRQONKIYKFRGTGKAWDEVDTENGNYDLYMADLMDHPEVINELR 249  
Db 251 NMGKVVNTNTIDGFRDLDAVKHIFKFFDMLSVYRSOTGKPLFTVGEYWSYDINKLHNY 310  
QY 71 TAVWIIPPAWKGTSONDVGYGAYLDLGEFNQKAVRTKYGTGRSQ-LQGAVTSLKNGIQ 129  
Db 132 VYADVVDHKGAGDTEWDAVEVNPDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 191  
QY 130 VYGDVVMNHKGAGDTEWDAVEVNPDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 189  
Db 192 WYHFGDWDDESKL-SRIYKFRGIGKAWDEVDTENGNYDLYMADLMDHPEVVTTELK 250  
QY 190 WYHFGDWDQSRQONKIYKFRGTGKAWDEVDTENGNYDLYMADLMDHPEVINELR 249  
Db 251 NMGKVVNTNTIDGFRDLDAVKHIFKFFDMLSVYRSOTGKPLFTVGEYWSYDINKLHNY 310  
QY 250 NMGVWYNTNLNDGFRIDAVKHISYTRDMLTHVRNTTGPMPFAVEAFKNDLAAIENY 309  
Db 311 ITKNTGMSLFDAPLHNNYTTASKSGGAFDMRTLMTNTLMKDOPTLAVTFVDNHDTPGQ 370  
QY 310 LNKTSWNHVSFVDFPLHYNLASNGGYEDMRNLNGSVVQKPHIAVTFVDNHDSPQGE 369  
Db 371 ALQSWDVPFKPLAYAFILTROEGYPCVFGYDYGIPQYNIPSLKSKIDPLLIARDYAY 430  
QY 370 ALESFVQSWFKPLAYALILTROEGYPSVFGYDYGTHGVFSMKSKIDPLLOARQTYAY 429  
Db 431 GTQHDYLDHSDIIGTWREGTEKPSGLAALITDGPGGSKMYGKHKHAGKYFYDLTGNR 490  
QY 430 GTQHDYFDHDDIIGTWREGDSSHPSGLATIMSDGPGKNMYGKHKAGQVWRDITGNR 489  
RESULT 4 PRELIMINARY; PRT; 549 AA.  
ID Q45519;  
AC Q45519;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE ALPHA AMYLASE PRECURSOR.  
OS BACILLUS STEAROTHERMOPHILUS.

Db 491 SDTITNSDGMGEKFGVNGSVVWV 515  
QY 490 SGTITNADGMGFTVNGGAVSVWV 514  
RESULT 3 PRELIMINARY; PRT; 549 AA.  
ID Q31193;  
AC Q31193;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE ALPHA AMYLASE.  
GN AMI.  
OS BACILLUS STEAROTHERMOPHILUS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 31195;  
RA DA SILVA A.C.R., FERNANDES E., PUEYO M.T.;  
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: AF032864; G2642326; -.  
SQ SEQUENCE 549 AA; 62651 MW; D9051082 CRC32;  
Query Match 68.2%; Score 2642; DB 9; Length 549;  
Best Local Similarity 67.1%; Pred. No. 0.00e+00;  
Matches 339; Conservative 79; Mismatches 82; Indels 5; Gaps 5;  
Db 15 LLAFLITAS-LFCPTGQPAKAA-AFENGTMQYFEWLPDDGTLTWKVANEANLSSIGI 72  
QY 11 LTLTLLAVAVLPFYMTPEQAHHNGTNGTMQYFEWHLPDNGNHNRLRDAANLKSIGI 70  
Db 73 TALMLPPAYKTSRSDVGYDYLDLGEFNQKAVRTKYGTKAQYLO-AIOAAHAGMQ 131  
QY 71 TAVWIIPPAWKGTSONDVGYGAYLDLGEFNQKAVRTKYGTGRSQ-LQGAVTSLKNGIQ 129  
Db 132 VYADVVDHKGAGDTEWDAVEVNPDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 191  
QY 130 VYGDVVMNHKGAGDTEWDAVEVNPDRNOEISGTQIOAWTKFDPGGRNTYSSFKWR 189  
Db 192 WYHFGDWDDESKL-SRIYKFRGIGKAWDEVDTENGNYDLYMADLMDHPEVVTTELK 250  
QY 190 WYHFGDWDQSRQONKIYKFRGTGKAWDEVDTENGNYDLYMADLMDHPEVINELR 249  
Db 251 NMGKVVNTNTIDGFRDLDAVKHIFKFFDMLSVYRSOTGKPLFTVGEYWSYDINKLHNY 310  
QY 250 NMGVWYNTNLNDGFRIDAVKHISYTRDMLTHVRNTTGPMPFAVEAFKNDLAAIENY 309  
Db 311 ITKNTGMSLFDAPLHNNYTTASKSGGAFDMRTLMTNTLMKDOPTLAVTFVDNHDTPGQ 370  
QY 310 LNKTSWNHVSFVDFPLHYNLASNGGYEDMRNLNGSVVQKPHIAVTFVDNHDSPQGE 369  
Db 371 ALQSWDVPFKPLAYAFILTROEGYPCVFGYDYGIPQYNIPSLKSKIDPLLIARDYAY 430  
QY 370 ALESFVQSWFKPLAYALILTROEGYPSVFGYDYGTHGVFSMKSKIDPLLOARQTYAY 429  
Db 431 GTQHDYLDHSDIIGTWREGTEKPSGLAALITDGPGGSKMYGKHKHAGKYFYDLTGNR 490  
QY 430 GTQHDYFDHDDIIGTWREGDSSHPSGLATIMSDGPGKNMYGKHKAGQVWRDITGNR 489  
RESULT 4 PRELIMINARY; PRT; 549 AA.  
ID Q45519;  
AC Q45519;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE ALPHA AMYLASE PRECURSOR.  
OS BACILLUS STEAROTHERMOPHILUS.



QY 36 TNGTMMQYFEWHL.PNDGNHNRRLRDAANLKSIGITAVVWIPPAWKGTSONDVGYGAYDLY 95  
 Db 62 DLGEFNQKGTVTKYGTADYLAISALKNNGIEPLADYILNHHKAAADHTETFKVVEVAP 121  
 QY 96 DLGEFNQKGTVTKYGTADYLAISALKNNGIEPLADYILNHHKAAADHTETFKVVEVAP 155  
 Db 122 EDRKTVLQPPFEIEGWTETFEVAIVPNDFECHWHYHTGTYD-VKGTGKTGIFQIGDN 180  
 QY 156 SNRNOEISGEYFIEAWTKFDFPGRNTHSNFKRWYHFDGTDQDSRLQNKIYKFRGTG 215  
 Db 181 KGWANODLVGNGYDYLMDLADLKPKHEVINKNYDNADWVETTVGKGFELDAIKHID 240  
 QY 216 KAW-DWE-VDIENGNYDLYMADIDMDHPEVINELRNGVWYTNLNDGFRIDAVKHK 273  
 Db 241 SFFMGNFTDMKTKSGNDYFVGFENGDEKSNNDYLASTDYRFDLVDVRLHQLFEASK 300  
 QY 274 YSYTRDWTHTVNTTKGPMFAVEFKWKNDAIENLYNKTSNHSHSVDFVPLHYNLYNAS 333  
 Db 301 AKETYDLRQIFEOTLVKKNHPSAVTFVDNHDHTQGOALESTIEEWFKAAYALILLRGTG 360  
 QY 334 SGGYFDMNILSGSVQKHPHNAVTFVDNHDHDSQGEALESFQSWFKPLAYALILTREQ 393  
 Db 361 LPCIFYGYGISOFOAESQTVTDKLIELRNKNAVYQEMDYDQANCIGTWCIGDDEH 420  
 QY 394 YPSVFGYGYGIPTH-GVPSMKSKIDPLQARQYAYGTQHDYDHDHDIIGWTREGDSSH 452  
 Db 421 P-TALAGLINNSKATSRMFGKWKACKLFTDALGNQAAHVOIDEGYGDYLVGEKSVSA 479  
 QY 453 PNLGATIMSDGPGKWMYGVKKGAGQVWRDITGNRSGTITINADGNGFTVNGGAVSV 512  
 Db 480 WI 481  
 QY 513 WV 514

RESULT 7  
 ID O33476 PRELIMINARY; PRT; 461 AA.  
 AC O33476;  
 DT 01-JAN-1998 (TREMREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
 DE ALPHA-AMYLASE (EC 3.2.1.1) (ALPHA-AMYLASE)  
 GN APKA.  
 OS PYROCOCUS SP.  
 OC ARCHAEBACTERIA; CRENARCHAEOTA; THERMOPROTEALES; THERMOCOCCACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-KOD1;  
 RA TACHIBANA Y., MENDEZ L., FUJIWARA S., TAKAGI M., IMANAKA T.;  
 RL J. FERMENT. BIOENG. 82:224-232(1996).  
 DR EMBL; D83793; D1021976; -.  
 KW SIGNAL.  
 FT SIGNAL.  
 FT CHAIN 1 26 POTENTIAL.  
 SQ SEQUENCE 461 AA; 52213 MW; DF6927EF CRC32;

Query Match 9.5%; Score 368; DB 9; Length 461;  
 Best Local Similarity 34.8%; Pred.No. 3.77e-49;  
 Matches 49; Conservative 41; Mismatches 49; Indels 2; Gaps 2;

Db 2 KKFVALLIMFFVSVMAVAQAASAKYSELEEGGVIMQAFYWDVPAGGIWDTIRSKIP 61  
 QY 5 NRIISVLLTLLAVAVLPYTPAQAHN-GTNGTMMQYFEWHL.PNDGNHNRRLRDA 63  
 Db 62 EWEYAGISAIWIPPAKMGAYSMGYDYPDFDLEGEYKQGTETFRFGSKQLINMINT 121  
 QY 64 NLKSGITAVWIPPAWKGTSON-DVGYGAYDLYDLGEFNQKGTVRTKYGTRSOLOGAVTS 122  
 Db 122 AHAYGKIVADIYVNHRRAGD 142  
 QY 123 LKNNGIQVIGDVVNMHKGAD 143

RESULT 8  
 ID O08452 PRELIMINARY; PRT; 460 AA.  
 AC O08452;  
 DT 01-JUL-1997 (TREMREL. 04, CREATED)  
 DT 01-JUL-1997 (TREMREL. 04, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)  
 DE ALPHA AMYLASE (EC 3.2.1.1) (ALPHA-AMYLASE) (1,4-ALPHA-D-GLUCAN  
 GLUCANOHYDROLASE) (GLYCOGENASE).  
 GN AMYA.  
 OS PYROCOCUS FURIOSUS.  
 OC ARCHAEBACTERIA; CRENARCHAEOTA; THERMOPROTEALES; THERMOCOCCACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM 3638;  
 RA DONG G., VIELLE C., SAVCHENKO A., ZEIKUS J.G.;  
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM3638;  
 RX MEDLINE; 97341170.  
 RA JORGENSEN S., VORGAS C.E., ANTRANIKIAN G.;  
 RL J. BIOL. CHEM. 272:16335-16342(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM3638;  
 RA JOERGENSEN S.T., VORGAS C.E., ANTRANIKIAN G.;  
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 DR EMBL; AF001268; G2183106; -.  
 DR EMBL; U96622; G2343083; -.  
 KW HYDROLASE; GLYCOSIDASE.  
 SQ SEQUENCE 460 AA; 52909 MW; EDEC186F CRC32;

Query Match 9.3%; Score 362; DB 9; Length 460;  
 Best Local Similarity 37.3%; Pred.No. 6.41e-48;  
 Matches 53; Conservative 36; Mismatches 51; Indels 2; Gaps 2;

Db 1 MNIKKPLTLLTLLFFVLASPVSAKYLEEGGVIMQAFYWDVPGGIWHDHRSKI 59  
 QY 3 LHNRIISVLLTLLAVAVLPYTPAQAHN-GTNGTMMQYFEWHL.PNDGNHNRRLRDA 62  
 Db 60 PEWEYAGISAIWIPPAKMGAYSMGYDYPDFDLEGEYKQGTETFRFGSKQLVRLIQ 119  
 QY 63 ANLKSIGITAVWIPPAWKGTSON-DVGYGAYDLYDLGEFNQKGTVRTKYGTRSOLOGAVT 121  
 Db 120 TAHAYGKIVADIYVNHRRAGD 141  
 QY 122 LKNNGIQVIGDVVNMHKGAD 143

RESULT 9  
 ID O60224 PRELIMINARY; PRT; 504 AA.  
 AC O60224;  
 DT 01-NOV-1996 (TREMREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)  
 DE ALPHA-AMYLASE (EC 3.2.1.1) (ALPHA-AMYLASE)  
 OS NATRONOCOCUS SP.  
 OC ARCHAEBACTERIA; EURYARCHAEOTA; HALOBACTERIALES; HALOBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AH-36;  
 RX MEDLINE; 94327485.  
 RA KOBAYASHI T., KANAI H., AONO R., HORIKOSHI K., KUDO T.;  
 RL J. BACTERIOL. 176:5131-5134(1994).  
 DR EMBL; D26510; G517137; -.  
 KW SIGNAL.  
 FT SIGNAL.  
 FT CHAIN 1 43 POTENTIAL.  
 SQ SEQUENCE 504 AA; 55696 MW; A5E843BA CRC32;

Query Match 4.4%; Score 172; DB 9; Length 504;

```

SQ SEQUENCE 423 AA; 46996 MW; 343ABA5 CRC32;

Query Match      3.8%; Score 148; DB 8; Length 423;
Best Local Similarity 34.7%; Pred. No. 2,02e-07;
Matches          35; Conservative 24; Mismatches 31; Indels 11; Gaps 6;

Db    22 LFQGFNWSKQGQWYNLSLIPDLANAGITHVWLPPSSHSVSP-Q-GYMPGRLYDL- 78
     : ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy   40 MMGYFEWHLPDGNHW-NRLRDAANLKSGITAVVIPPAWKTSQNDGVGAYDYLDLG 98

Db    79 D-----A--SKYGNKOOQLSVALHAKGLKAVADIIVINHR 112
     : : |||| : |||| : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy   99 EFNQKGTVRTKYTRSOLOGAVTSLKNNGIQVYGDVVMMHK 139


RESULT 12
ID Q41442 PRELIMINARY; PRT; 407 AA.
AC Q41442;
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1996 (TREMBREL. 01, LAST ANNOTATION UPDATE)
DE ALPHA-AMYLASE PRECURSOR.
OS SOLANUM TUBEROSUM (POTATO).
OC EUKARYOTA; PLANTA; EMERYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC SOLANALES; SOLANACEAE.
RN [1]
RC STRAIN-DIANELLA; TISSUE=SPROUT;
RA GAUSING K., KREIBERG T.D.;
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR ENBL; M79328; G1220453; -
DR DR ENBL; A21341; G512394; -.
KW SIGNAL.
FT SIGNAL.
FT CHAIN 1 18 POTENTIAL.
FT FT CHAIN 19 407 ALPHA-AMYLAZE.
SQ SEQUENCE 407 AA; 46346 MW; B51E7F71 CRC32;

Query Match      3.5%; Score 135; DB 8; Length 407;
Best Local Similarity 21.8%; Pred. No. 2,27e-05;
Matches          39; Conservative 51; Mismatches 77; Indels 12; Gaps 10;

Db    155 PNIDHTQSFRVKDLIDMWRMRSSVGQDFRFDAKGYASKYKEYIEGAEPFAVGEYW 214
     : ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy   235 ADICMDHPHEINELRWNVYTNTLNLDGRIDAVKHIXSYTRDWLTTHVRNTTGKPMF- 293

Db    215 DTCNYGSNLDDYNODSHQRRIIWIWDGACOLSTAFETTKAVLOEAVKEFWLRDSKK 274
     : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy   294 AVAEFWKNDIA-AIENYLNT-SWNHS--FDYPLHY--N-LYNASNSGGYFDMRIL-N 345

Db    275 PGVLGLUPSRATFFIDHNIGTGOAHWP-FPSRHMEGYAVILTH-PGIPSVEFDHFY 331
     : : | : |||| : |||| : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy   346 GS-VVQKPHPTATFVDNHDSQFGEALESFVSQWFPLAYALILTREQGPSVFYGDYY 403


RESULT 13
ID QO3651 PRELIMINARY; PRT; 427 AA.
AC QO3651;
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DE 01-FEB-1997 (TREMBREL. 02, LAST ANNOTATION UPDATE)
DE ALPHA-AMYLASE TYPE B PRECURSOR.
GN AMY46.
OS HORDEUM VULGARE (BARLEY).
OC EUKARYOTA; PLANTA; EMERYOPHYTA; ANGIOSPERMAE; MONOCOTyledONEAE;
OC CYPERALES; GRAMINEAE.
RN [1]
RC KHURSHED B., ROGERS J.;
RA KHURSHED B., ROGERS J.;
RL J. BIOL. CHEM. 0:0-0(0);
DR ENBL; J04202; G166985; -.
KW SIGNAL.
FT SIGNAL.
FT FT CHAIN 1 24 POTENTIAL.
FT FT CHAIN 25 427 ALPHA-AMYLAZE.
```

SQ SEQUENCE 427 AA; 47456 MW; 2E2F49DA CRC32;

Query Match 3.5%; Score 135; DB 8; Length 427;  
Best Local Similarity 29.7%; Pred. No. 2.27e-05;  
Matches 30; Conservative 27; Mismatches 33; Indels 11; Gaps 6;

Db 27 LFOGFNWSKHWGTFNLMGKVDVDDIAAGVTHVWLPSPSHSVS-NE-GYMPGRLYDI- 83  
QY 40 MMQYFEWHLPNDDGNHW-NRLRDDAANKSKGITAVWIPPAWKTSQNDVGYGAYDLYDLG 98  
Db 84 D-----A--SKYGNKAOLKSLIGALHGKGVKAIADIVINHR 117  
QY 99 EFNQKGTVRTKYGTRSLQGAVTSLKNGIQVYGDVVMNHK 139

RESULT 14  
ID Q41441 PRELIMINARY; PRT; 349 AA.  
AC Q41441;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DE ALPHA-AMYLASE (FRAGMENT).  
GN AMY23.  
OS SOLANUM TUBEROSUM (POTATO).  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOITYLEDONEAE;  
OC SOLANALES; SOLANACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DIANELLA; TISSUE=SPROUT;  
RA GAUSING K., KREIBERG T.D.;  
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; M81682; G1220451;  
DR EMBL; A21345; G579229;  
FT NON\_TER 1  
SQ SEQUENCE 349 AA; 39294 MW; 64DC2187 CRC32;

Query Match 3.4%; Score 130; DB 8; Length 349;  
Best Local Similarity 41.2%; Pred. No. 1.31e-04;  
Matches 21; Conservative 11; Mismatches 17; Indels 2; Gaps 2;

Db 223 PKRAVTFIDNHDTGTONMPPPSDKVMQ-GYAYILTH-PGIPSVFYDHF 271  
QY 353 PIHATFVDNHDSPGEALESFVQSWFKPLAYALILTREQGYPSVYGDY 403

RESULT 15  
ID Q42504 PRELIMINARY; PRT; 421 AA.  
AC Q42504;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DE ALPHA-AMYLASE TYPE A.  
GN AMY32B.  
OS HORDEUM VULGARE (BARLEY).  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; MONOCOTYLEDONEAE;  
OC CYPERALES; GRAMINEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA WHITTIER R.F., DEAN D.A., ROGERS J.C.;  
RL NUCLEIC ACIDS RES. 13:0-0(1987).  
RN [2]  
RP SEQUENCE OF 1-46 FROM N.A.  
RX MEDLINE; 85006965.  
RA ROGERS J.C., MILLIMAN C.;  
RL J. BIOL. CHEM. 259:12234-12240(1984).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA WHITTIER R.F., DEAN D.A., ROGERS J.C.;  
RL SUBMITTED (FEB-1991) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; X05166; G19049;  
DR EMBL; M15208; G167001;  
SQ SEQUENCE 421 AA; 46102 MW; EAB3DF68 CRC32;

Query Match 3.4%; Score 132; DB 8; Length 421;  
Best Local Similarity 28.7%; Pred. No. 6.53e-05;  
Matches 29; Conservative 30; Mismatches 31; Indels 11; Gaps 6;

Db 27 LFOGFNWSKHWGTFNLMGKVDVDDIAAGVTHVWLPSPSHSVS-NE-GYMPGRLYDI- 83  
QY 40 MMQYFEWHLPNDDGNHW-NRLRDDAANKSKGITAVWIPPAWKTSQNDVGYGAYDLYDLG 98  
Db 84 D-----A--SKYGNKAOLKSLIGALHGKGVKAIADIVINHR 117  
QY 99 EFNQKGTVRTKYGTRSLQGAVTSLKNGIQVYGDVVMNHK 139

Search completed: Tue Sep 15 14:21:46 1998  
Job time : 131 secs.

T)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

| Result | Query |       |       | DB | ID     | Description           | Pred. No. |
|--------|-------|-------|-------|----|--------|-----------------------|-----------|
|        | No.   | Score | Match |    |        |                       |           |
| 1      | 3368  | 87.0  | 518   | 1  | A27705 | alpha-amylase (EC 3.2 | 0.00e+00  |
| 2      | 2679  | 69.2  | 512   | 1  | AB5SL  | alpha-amylase (EC 3.2 | 0.00e+00  |
| 3      | 2661  | 68.7  | 549   | 1  | A3541  | alpha-amylase (EC 3.2 | 0.00e+00  |
| 4      | 2645  | 68.3  | 549   | 1  | A24436 | alpha-amylase (EC 3.2 | 0.00e+00  |
| 5      | 2623  | 67.7  | 549   | 1  | A24549 | alpha-amylase (EC 3.2 | 0.00e+00  |
| 6      | 2559  | 66.1  | 514   | 1  | ALBSN  | alpha-amylase (EC 3.2 | 0.00e+00  |
| 7      | 2524  | 65.2  | 548   | 1  | ALBSF  | alpha-amylase (EC 3.2 | 0.00e+00  |
| 8      | 1912  | 49.4  | 493   | 2  | S15713 | alpha-amylase (EC 3.2 | 0.00e+00  |
| 9      | 1558  | 40.2  | 290   | 5  | 1BPLB  | alpha-1,4-glucan-4-gl | 1.00e-25  |
| 10     | 1558  | 40.2  | 290   | 5  | 1VJ52  | alpha-amylase (EC 3.2 | 1.00e-25  |
| 11     | 1403  | 36.2  | 494   | 1  | B45738 | alpha-amylase (EC 3.2 | 2.00e-23  |
| 12     | 1360  | 35.1  | 495   | 1  | A45738 | alpha-amylase (EC 3.2 | 2.62e-22  |
| 13     | 1006  | 26.0  | 179   | 5  | 1BPLA  | alpha-1,4-glucan-4-gl | 6.06e-15  |
| 14     | 1006  | 26.0  | 179   | 5  | 1VJ51  | alpha-amylase (EC 3.2 | 6.06e-15  |
| 15     | 344   | 8.9   | 217   | 2  | A19506 | alpha-amylase (EC 3.2 | 1.87e-36  |
| 16     | 217   | 5.6   | 547   | 2  | A32803 | glucan 1,4-alpha-malt | 7.94e-16  |
| 17     | 218   | 5.6   | 551   | 2  | S05667 | glucan 1,4-alpha-malt | 5.59e-16  |
| 18     | 203   | 5.2   | 417   | 5  | 1AMG   | 1,4-alpha-D-glucan ma | 1.08e-13  |
| 19     | 203   | 5.2   | 418   | 5  | 1YDD   | 1,4-alpha maltotetra  | 1.08e-13  |
| 20     | 203   | 5.2   | 418   | 5  | 1YDC   | 1,4-alpha maltotetra  | 1.08e-13  |
| 21     | 203   | 5.2   | 418   | 5  | 1YDA   | 1,4-alpha maltotetra  | 1.08e-13  |
| 22     | 202   | 5.2   | 528   | 1  | ALBSK  | alpha-amylase (EC 3.2 | 1.52e-13  |
| 23     | 192   | 5.0   | 713   | 2  | S09196 | alpha-amylase (EC 3.2 | 4.77e-12  |

```

Db 62 SDASNLSKSGITAVIPPAWKGASQNDVGYGAYDLYDLGBFNOKGTVRTYKGYTRSQLOAA 121
QY 60 DDAANLSKSGITAVIPPAWKGASQNDVGYGAYDLYDLGBFNOKGTVRTYKGYTRSQLOQA 119
Db 122 VTSLKNGIQVYGDVYNNHKGADATMRAVEVNNRNNOEYTGVTIEAWTRFDPGR 181
QY 120 VTSLKNGIQVYGDVYNNHKGADATMRAVEVNNRNNOEYTGVTIEAWTRFDPGR 179
Db 182 GNTHSSFKRWYHFDGVDWDSRLNRIYKFRGHGKAWDEVDTEGNTDYLMIADIDM 241
QY 180 GNTHSFKRWYHFDGVDWDSRLNRIYKFRGHGKAWDEVDTEGNTDYLMIADIDM 239
Db 242 DRPEVYNELRNWGWYNTNLGLGDFRIDAVKHIIKYSTRWINHVSATGNMFVAEWF 301
QY 240 DRPEVYNELRNWGWYNTNLGLGDFRIDAVKHIIKYSTRWINHVSATGNMFVAEWF 299
Db 302 KNDLGAENYLOKTNHNSVEDYPLHNLNASKSGNYDMRINFNGTVQVRPHSHAVTF 361
QY 300 KNDLGAENYLOKTNHNSVEDYPLHNLNASKSGNYDMRINFNGTVQVRPHSHAVTF 359
Db 362 VDNHDSQPEALESFVQSNFKPLAYALLTREGQYPSVFGDYIGYPTGHVPMARSKIDP 421
QY 360 VDNHDSQPEALESFVQSNFKPLAYALLTREGQYPSVFGDYIGYPTGHVPMARSKIDP 419
Db 422 ILEARKYAYKQNDYLDHNNIIGTWREGTAHPNGLATIMSDGAGGSKWFMVGRNKAG 481
QY 420 ILEARKYAYKQNDYLDHNNIIGTWREGTAHPNGLATIMSDGAGGSKWFMVGRNKAG 479
Db 482 QVMSDITGNRTGVTINADGWNFSYNGSGSVIWN 517
QY 480 QVMSDITGNRTGVTINADGWNFSYNGSGSVIWN 515

RESULT 2
ENTRY ALBLSL #type complete
TITLE alpha-amylase (EC 3.2.1.1) precursor - Bacillus licheniformis
ALTERNATE_NAMES 1,4-alpha-D-glucan glucanohydrolase
ORGANISM #formal_name Bacillus licheniformis
DATE 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 20-Mar-1998
ACCESSION A91997; B24549; A91796; A21663; A26151; S53788; I39772;
I39774; A00844
REFERENCE A91997
#authors Yuuki, T.; Nomura, T.; Tezuka, H.; Tsuboi, A.; Yamagata, H.;
Tsukagoshi, N.; Uda, S.
#journal J. Biochem. (1985) 98:1147-1156
#title Complete nucleotide sequence of a gene coding for heat- and
pH-stable alpha-amylase of Bacillus licheniformis:
Comparison of the amino acid sequences of three bacterial
liquefying alpha-amylases deduced from the DNA sequences.
#cross-references MUID:86111694
#accession A91997
#molecule_type DNA
#residues 1-512 #label YUU
#cross-references GB:X03236; NID:g39551; PID:g39552
#experimental_source ATCC 27811
REFERENCE A91817
#authors Gray, G.L.; Mainzer, S.E.; Rey, M.W.; Lamsa, M.H.; Kindle,
K.L.; Carmona, C.; Requa, C.
#journal J. Bacteriol. (1986) 166:635-643
#title Structural genes encoding the thermophilic alpha-amylases of
Bacillus stearothermophilus and Bacillus licheniformis.
#cross-references MUID:86195857
#accession B24549
#molecule_type DNA
#residues 1-162,'L',164-338,'G',340-348,'S',350-512 #label GRA
#cross-references GB:M13256; NID:g142510; PID:g142511
#experimental_source NCIB 8061
REFERENCE A91796
#authors Stephens, M.A.; Ortlepp, S.A.; Ollington, J.F.; McConnell,
D.J.
#journal J. Bacteriol. (1984) 158:369-372
#title Nucleotide sequence of the 5' region of the Bacillus

```

```

licheniformis alpha-amylase gene: comparison with the
Bacillus amyloliquefaciens gene.
#cross-references MUID:84185455
#accession A91796
#molecule_type DNA
#residues 1-104 #label STE
#cross-references GB:K01984; NID:g142432; PID:g142433
REFERENCE A21663
#authors Sibakov, M.; Palva, I.
#journal Eur. J. Biochem. (1984) 145:567-572
#title Isolation and the 5'-end nucleotide sequence of Bacillus
licheniformis alpha-amylase gene.
#cross-references MUID:85076654
#accession A21663
#molecule_type DNA
#residues 1-3,'H',5-12,'P',14-47,'R',49-61,'V',63,'D',65-67,'VA',
70-71,'S',73-80,'D',82-104,118-121 #label STB
#experimental_source chromosomal DNA of ATCC 14580
#note the authors translated the codon CGT for residue 48 as
Gly and GAC for residue 64 as His
REFERENCE A26151
#authors Kuhn, H.; Fietzek, P.; Lampen, J.O.
#journal J. Bacteriol. (1982) 149:372-373
#title N-terminal amino acid sequence of Bacillus licheniformis
alpha-amylase: comparison with Bacillus amyloliquefaciens
and Bacillus subtilis enzymes.
#cross-references MUID:82098050
#accession A26151
#molecule_type protein
#residues 30-37,'E',39-41,'X',43-47 #label KUH
REFERENCE S53788
#authors Machius, M.; Wiegand, G.; Huber, R.
#journal J. Mol. Biol. (1995) 246:545-559
#title Crystal structure of calcium-depleted Bacillus licheniformis
alpha-amylase at 2.2 A resolution.
#accession S53788
#molecule_type protein
#residues 1-220-227 #label MAC
#note sequence represents amino end of an internal fragment
created by a single enzymatic cleavage by a protease
trace contaminant during purification
REFERENCE I39772
#authors Jorgensen, P.L.; Hansen, C.K.; Poulsen, G.B.; Diderichsen, B.
#journal Gene (1990) 96:37-41
#title In vivo genetic engineering: homologous recombination as a
tool for plasmid construction.
#cross-references MUID:91092499
#accession I39772
#status translated from GB/EMBL/DBDJ
#molecule_type DNA
#residues 1-32,'I' #label RES
#cross-references GB:M62637; NID:g142498; PID:g142499
REFERENCE I39773
#authors Laolde, B.M.; Chambliss, G.H.; McConnell, D.J.
#journal J. Bacteriol. (1989) 171:2435-2442
#title Bacillus licheniformis alpha-amylase gene, amyL, is subject
to promoter-independent catabolite repression in Bacillus
subtilis
#cross-references MUID:89213924
#accession I39774
#status translated from GB/EMBL/DBDJ
#molecule_type DNA
#residues 1-29 #label RE2
#cross-references GB:M26412; NID:g341477; PID:g516590
GENETICS amyL
#gene catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic
bonds
#pathway glycogen/starch degradation
#CLASSIFICATION #superfamily alpha-amylase, amyloliquefaciens type;
alpha-amylase core homology
#KEYWORDS extracellular protein; glycosidase; heat-stable protein;

```





```
##residues 1-549 #label NAK
##cross-references GB:M1450
##experimental_source plasmid pAT5
##note
amino end of the mature protein also determined
REFERENCE
I39772
Jorgensen, P.L.; Hansen, C.K.; Poulsen, G.B.; Diderichsen, B.
Gene (1990) 96:37-41
In vivo genetic engineering: homologous recombination as a
tool for plasmid construction.
#cross-references MUID:91092499
#accession I39777
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-45 #label RES
#cross-references GB:M62638; NID:g142514; PID:g142515
COMMENT
Alpha-amylase genes have been found on plasmids and in multiple
copies on the chromosome in various strains of this organism.
GENETICS
#gene amyS
#genome plasmid
#start_codon GTG
FUNCTION
#description catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic
bonds
#pathway glycogen/starch degradation
CLASSIFICATION
#superfamily alpha-amylase, amylioliquefaciens type;
alpha-amylase core homology
KEYWORDS
extracellular protein; glycosidase; heat-stable protein;
hydrolase; polysaccharide degradation
FEATURE
1-34
#domain signal sequence #status predicted #label SIG\
35-549
#product alpha-amylase #status experimental #label MAT\
235-368
#domain alpha-amylase core homology #label AMY\
139,237,272
#binding_site calcium (Asp, Asp, His) #status predicted\
268,298,365
#active_site Asp, Glu, Asp #status predicted
SUMMARY
#length 549 #molecular-weight 62670 #checksum 5048
Query Match 68.3%; Score 2645; DB 1; Length 549;
Best Local Similarity 67.1%; Pred. No. 0.00e+00;
Matches 339; Conservative 81; Mismatches 80; Indels 5; Gaps 5;
Db 15 LIAFLLT-ALLFPCPTGCPAKAA-APFNGTMMQYFEWYLPDDGTLTKVANEANNLSSIGI 72
Qy 11 LTLTLLAVLFPYTPTEPAQAHHNGTNGTMMQYFEWHLPNDRHNRLLRDAANLKSIGI 70
Db 73 TALWLPAYKGRSDVGVYDLDLGFEFNOKGAVRKYGTAKYLO-AIQAAHAAGMQ 131
Qy 71 TAVWIPPAWKGTQNDVGYGAYDLGFEFNOKGAVRKYGTGRSQ-LOGAVTSLKNNIGQ 129
Db 132 VYADVDFDHKGADGETWYDAVEVNPDRNQISGTGQIQAQTKDFPGRGNTYSFQKWR 191
Qy 130 VYGDVVMNHKGADGETWYNAVEVNRNRNQISGEYITEATKDFPGRGNTHSFKWR 189
Db 192 WYHFDGVDWDESKL-SRIYKFRGIGKAWDVEVDENGNYDLYMLADLMDHPEVVTEUK 250
Qy 190 WYHFDGTDWQSKLQNKIYKFRGTGKAWDVEVDIENGNYDLYMLADLMDHPEVINELR 249
Db 251 SNGKVVYNTNIDGFRDLDAVKIKFSFFPDWLSDFRSOTGKPLFTVGYEWSYDINKLHNY 310
Qy 250 NMGVYNTNLDGFRDAVDAVKHKYSYTRDLWTHVRNTTGKMFVFAEFAEKKNDLAAIENY 309
Db 311 INKNTGNTMSLFDAPLHNKYTASKSGTGMRTLTMTNLTKMDQPTLAVTFVDNHDPEPGQ 370
Qy 310 LNKTSNHSVDFVPLHNLNYSNASSGDFDMRNILNGSVVQKHPIHAVTFVDNHDSPQGE 369
Db 371 ALOSVDVDFPKPLAYAFILTRQEGYPCVYGYGIGIPQYNIPSLKSIDPLLIARDYAY 430
Qy 370 ALESFQSVFKPLAYALILTRQEGYPCVYGYGIGIPTHGVPSMKSIDPLLIQAROTYAY 429
Db 431 GTQHDYLDHSDIIGWTRREGVTEKPGSGLAALITDGPGGSKWYVGVQKHAGKVFDYLTGNR 490
Qy 430 GTQHDYDFHHDIIIGWTRREGDSSHPNSSLATIMSDGGGKNKMYVGVKHKAGQVWRDITGNR 489
```

```
Db 491 SDVTINSDGWBEKFGVNGSVSVVWV 515
Qy 490 SGTIINADGWNGFTVNGGAVSVVWV 514

RESULT 5
ENTRY A24549 #type complete
TITLE alpha-amylase (EC 3.2.1.1) precursor - Bacillus
stearothermophilus (strain NZ-3)
ALTERNATE_NAMES 1,4-alpha-D-glucan glucanohydrolase
ORGANISM #formal_name Bacillus stearothermophilus
DATE 30-Jun-1988 #sequence_revision 18-Aug-1995 #text_change
05-Sep-1997
ACCESSIONS A24549; I39501; I39770
REFERENCE A91817
#authors Gray, G.L.; Mainzer, S.E.; Rey, M.W.; Lamsa, M.H.; Kindle,
J. Bacteriol. (1986) 166:633-643
#journal J. Bacteriol. (1986) 166:633-643
#title Structural genes encoding the thermophilic alpha-amylases of
Bacillus stearothermophilus and Bacillus licheniformis.
#cross-references MUID:86195857
#accession A24549
#molecule_type DNA
#residues 1-549 #label GRA
#cross-references GB:M13255; NID:g142512; PID:g142513
#experimental_source genomic DNA of strain NZ-3
REFERENCE I39501
#authors Satoh, H.; Nishida, H.; Isono, K.
#journal J. Bacteriol. (1988) 170:1034-1040
#title Evidence for movement of the alpha-amylase gene into two
phylogenetically distant Bacillus stearothermophilus
strains.
#cross-references MUID:88139156
#accession I39501
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 536-549 #label RES
#cross-references GB:M29577; NID:g142476; PID:g142478
#experimental_source strain DY-5
#accession I39770
#status translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 536-549 #label RE2
#cross-references GB:M29578; NID:g142484; PID:g142486
#experimental_source strain 799
COMMENT Alpha-amylase genes have been found on plasmids and in multiple
copies on the chromosome in various strains of this organism.
GENETICS
#start_codon GTG
FUNCTION
#description catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic
bonds
#pathway glycogen/starch degradation
CLASSIFICATION
#superfamily alpha-amylase, amylioliquefaciens type;
alpha-amylase core homology
KEYWORDS
extracellular protein; glycosidase; heat-stable protein;
hydrolase; polysaccharide degradation
FEATURE
1-34
#domain signal sequence #status predicted #label SIG\
35-549
#product alpha-amylase #status predicted #label MAT\
235-368
#domain alpha-amylase core homology #label AMY\
139,237,272
#binding_site calcium (Asp, Asp, His) #status predicted\
268,298,365
#active_site Asp, Glu, Asp #status predicted
SUMMARY
#length 549 #molecular-weight 62643 #checksum 6769
Query Match 67.7%; Score 2623; DB 1; Length 549;
Best Local Similarity 66.9%; Pred. No. 0.00e+00;
Matches 338; Conservative 79; Mismatches 83; Indels 5; Gaps 5;
Db 15 LIAFWLTAS-LFCPTGCPAKAA-APFNGTMMQYFEWYLPDDGTLTKVANEANNLSSIGI 72
Qy 11 LTLTLLAVLFPYTPTEPAQAHHNGTNGTMMQYFEWHLPNDRHNRLLRDAANLKSIGI 70
```

```

Db 73 TALWLPAYKTSRSDYGYGYDLYDLGEFNQKGTVRTKTKYQKQYQ-ATQAAHAGMQ 131
QY 71 TAVWIPPAWKTSNDYGYGYDLYDLGEFNQKGTVRTKTKYQKQYQ-LOQAVTSLKNNIGQ 129
Db 132 VYADVDFDHKGAGDTEWDAVEVNPDSRNOEISGTIOQAWTKDFDNGRNTYSSFKWR 191
QY 130 VYGDVVMNHRGGADGTENVNAVENRNRNOEISGEYITAEWTKDFDNGRNTYSSFKWR 189
Db 192 WYHFGDGVDMDESRL-SRIYKFRGIGKAWDWEVDTEGNYDLYADLMDHPHVVTELK 250
QY 190 WYHFGDGTDMDSRQLNKIYKFRGTGKAWDWEVDTEGNYDLYADLMDHPHVVTELK 249
Db 251 NWGKYYNTNIDGFRDLDAVHKHIFSPDPLSVRSOTGAPLFTVGYGYSYDINKLNY 310
QY 250 NWGYYNTNLDGFRDLDAVHKHIIYTRDNLTHVRNTGKPMFAVEAFWKNDLAAIENY 309
Db 311 ITKNTGMSLFDAPLHNKFFYASKSGGAFDMSTLMNNTLMKDQPTLAVTEVDNHDTEPGQ 370
QY 310 LNKTSWNHVSFVDPVPLHYNLNASNGGYFDMRNILNGSVVQKPHAVTFVDNHDSDQGE 369
Db 371 ALOSVDPMFKPLAYAFILTRQEGYPCVFGYGYGIPQINIPSLKSKIDPLLIARDYAY 430
QY 370 ALESFVQSWFKPLAYALILTRQEGYPCVFGYGYGIPTHGVPSMKSKIDPLLIARQYAY 429
Db 431 GTQHDYLDHSDIIGWTRGCVTEKPGSLAALITDGPCKGKMYVQKQHAGKVFDLTGNR 490
QY 430 GTQHDYFDHDDIIGWTRGDSHNSGLATIMSDGPGGNKMYVQKQHAGKVFDLTGNR 489
Db 491 SDVTYINSDGWGEFKVNGSGSVVW 515
QY 490 SGTVTIIRADGNGTNGGAVSVW 514

RESULT 6
ENTRY #type complete
TITLE alpha-amylase (EC 3.2.1.1) precursor - Bacillus
ALTERNATE_NAMES amylioliquefaciens
ORGANISM 1,4-alpha-D-glucan glucohydrolase
#formal_name Bacillus amylioliquefaciens
DATE 30-Nov-1980 #sequence_revision 30-Jun-1987 #text_change
20-Mar-1998
ACCESSIONS A92389; A90307; I39756; I39763; A00843
REFERENCE A92389
#authors Takkinen, K.; Pettersson, R.F.; Kalkkinen, N.; Palva, I.;
Soderlund, H.; Kaariainen, L.
#journal J. Biol. Chem. (1983) 258:1007-1013
#title Amino acid sequence of alpha-amylase from Bacillus
amylioliquefaciens deduced from the nucleotide sequence of
the cloned gene.
#cross-references MUID:83108808
#contents PUB110
#accession A92389
#molecule_type DNA
#residues 1-514 #label TAK
#cross-references GB:J01542; GB:J01543; GB:M12033; GB:M12034;
NID:g142428; PID:g142429
REFERENCE A90307
#authors Chung, H.S.; Friedberg, F.
#journal Biochem. J. (1980) 185:387-395
#title Sequence of the N-terminal half of Bacillus amylioliquefaciens
alpha-amylase.
#cross-references MUID:80241725
#accession A90307
#molecule_type protein
#residues 32-553, 'I', 55-63, 'L', 65-78, 'D', 80-83, 'S', 85-222 #label
CHU
REFERENCE I39756
#authors Palva, I.; Pettersson, R.F.; Kalkkinen, N.; Lehtovaara, P.;
Sarvas, M.; Soderlund, H.; Takkinen, K.; Kaariainen, L.
#journal Gene (1981) 15:43-51
#title Nucleotide sequence of the promoter and NH2-terminal signal
peptide region of the alpha-amylase gene from Bacillus
amylioliquefaciens.

```

```

#cross-references MUID:82051296
#accession I39756
#status translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-96 #label RES
#cross-references EMBL:V00092; NID:g39297; PID:g39298
REFERENCE I39763
#authors Ruohonen, L.; Hackman, P.; Lehtovaara, P.; Knowles, J.K.C.;
Karaenen, S.
#journal Gene (1987) 59:161-170
#title Efficient secretion of Bacillus amylioliquefaciens
alpha-amylase cells by its own signal peptide from
Saccharomyces cerevisiae host.
#cross-references MUID:88137952
#accession I39763
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-39 #label RE2
#cross-references GB:M18424; NID:g142430; PID:g142431
FUNCTION
#description catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic
bonds
#pathway glycogen/starch degradation
#CLASSIFICATION #superfamily alpha-amylase, amylioliquefaciens type;
alpha-amylase core homology
KEYWORDS extracellular protein; glycosidase; hydrolase; polysaccharide
degradation
FEATURE
1-31 #domain signal sequence #status predicted #label SIG\
32-514 #product alpha-amylase #status predicted #label MPT\
229-362 #domain alpha-amylase core homology #label AMY\
133,231,266 #binding_site calcium (Asn, His) #status predicted\
262,292,359 #active_site Asp, Glu, Asp #status predicted
SUMMARY #length 514 #molecular-weight 58403 #checksum 2384
Query Match 66.1%; Score 2559; DB 1; Length 514;
Best Local Similarity 64.0%; Pred. No. 0.00e+00;
Matches 330; Conservative 90; Mismatches 89; Indels 7; Gaps 5;
Db 1 MIQKRKRTVSFRLVLMCTLLF--VSLPI-TKTSAVNGTLMQYFEWYTPNDGQHWKRLQND 57
QY 2 KLHNRIISVLLTLLLAVALVFPYMTPEQAHHNGTMMQYFEWHLFNDGNHNRURDD 61
Db 58 AEHLSDIGITAVWIPPAKGLSQSDNGYGPYDLYDLGEFOOKGTVRTKTKYQKQYQ-ATQAA 117
QY 62 ANLASKGITAVWIPPAWAGTSONDVGYGAYDLYDLGEFNQKGTVRTKTKYQKQYQ-ATQAA 121
Db 118 SLHSRNVQYGVYDVLNHNKAGADATEDVTAVERNPNRNOETSEYQIKAWTKDFRPGRN 177
QY 122 SLKNNGIQYGVYDVLNHNKAGADGTENVNAVENRNRNOEISGEYITAEWTKDFRPGRN 181
Db 178 TYSDFKWHYHFDGADWDESRI-SRIFKRGKAWDWEVSENGNYDLYADVDYD 236
QY 182 THSNFKRWYHFDGTDWQSRQLNKIYKFRGTGKAWDWEVDTEGNYDLYADIDMDH 241
Db 237 PDVVAETKKGWIYANELSDIGFRIDAARKHIFSLRDVQAVQATGKEMFTVAEYQW 296
QY 242 PEVIELRWGYYNTNLDGFRIDAARKHIIYTRDNLTHVRNTGKPMFAVEAFWKN 301
Db 297 NAGKLENYLNTKTSFNQSVDFVPLHFNLAQASQGGYDMRRLDGTVYVSRHPKAVTFVE 356
QY 302 DLAAIENYLNKTSWNHVSFVDPVPLHYNLNASNGGYFDMRNILNGSVVQKPHAVTFVD 361
Db 357 NHDTPQGSLSTVQWTFKPLAYAFILTRSGYPCVFGYGYGIPQINIPSLKSKIDPLLIAR 416
QY 362 NHDSPGEALSFVQSWFKPLAYALILTRQEGYPCVFGYGYGIPTHGVPSMKSKID 418
Db 417 PILKARKEVAYQPHDYIDHPDVIGWTRGDSHNSGLATIMSDGPGGNKMYVQKQHAGK 476
QY 419 PLLQARQTYAYGTQHDYFDHDDIIGWTRGDSHNSGLATIMSDGPGGNKMYVQKQHAGK 478
Db 477 GETWYDITGNRSYTVKIGSDGGEFHVNDGVSIVV 512

```

|                       |  |   |                                      |
|-----------------------|--|---|--------------------------------------|
| QY                    | 479  | QGVWRDITGNRSCTVTINADGNGFTVNGGAVSVWV                       | 514                                  |
| ENTRY                 | 7  |   |                                      |
| TITLE                 | ALBSF  | #type complete  |                                      |
| ALTERNATE_NAMES       | alpha-amylase (EC 3.2.1.1)   | precursor - Bacillus                                      |                                      |
| ORGANISM              | stearothermophilus (strain DY-5)   | plasmid PHI300  |                                      |
| DATE                  | 1,4-alpha-D-glucan glucanohydrolase  |   |                                      |
|                       | #formal_name Bacillus stearothermophilus   |   |                                      |
|                       | 30-Jun-1987 #sequence_revision 30-Jun-1987   | #text_change  |                                      |
|                       | 16-Feb-1997  |   |                                      |
| ACCESSIONS            | A91999; B91999; A91804; A00845   |   |                                      |
| REFERENCE             | A91999   |   |                                      |
| #authors              | Ihara, H.; Sasaki, T.; Tsuboi, A.; Yamagata, H.; Tsukagoshi, N.; Uda, S.   |   |                                      |
| #journal              | J. Biochem. (1985) 98:95-103   |   |                                      |
| #title                | Complete nucleotide sequence of a thermophilic alpha-amylase gene: homology between prokaryotic and eukaryotic alpha-amylases at the active sites.               |   |                                      |
| #cross-references     | MUID:86008166  |   |                                      |
| #accession            | A91999   |   |                                      |
| #molecule_type        | DNA  |   |                                      |
| #residues             | 1-548  | #label IH1  |                                      |
| #cross-references     | GB:X02769  |   |                                      |
| #experimental_source  | plasmid PHI300   | from strain DY-5  |                                      |
| #accession            | B91999   |   |                                      |
| #molecule_type        | protein  |   |                                      |
| #residues             | 35-48  | #label IH2  |                                      |
| #experimental_source  | strain DY-5  |   |                                      |
| REFERENCE             | A91804   |   |                                      |
| #authors              | Tsukagoshi, N.; Iritani, S.; Sasaki, T.; Takemura, T.; Ihara, H.; Idota, Y.; Yamagata, H.; Uda, S.   |   |                                      |
| #journal              | J. Bacteriol. (1985) 164:1182-1187   |   |                                      |
| #title                | Efficient synthesis and secretion of a thermophilic alpha-amylase by protein-producing Bacillus brevis 47 carrying the Bacillus stearothermophilus amylase gene. |   |                                      |
| #cross-references     | MUID:86059211  |   |                                      |
| #contents             | pbAM101  |   |                                      |
| #accession            | A91804   |   |                                      |
| #molecule_type        | DNA  |   |                                      |
| #residues             | 1-29, 'O', '31-75, 'W', '77-122  | #label TSU  |                                      |
| COMMENT               | Alpha-amylase genes have been found on plasmids and in multiple copies on the chromosome in various strains of this organism.                                    |   |                                      |
| GENETICS              |  |   |                                      |
| #genome               | plasmid  |   |                                      |
| #start_codon          | GTC  |   |                                      |
| FUNCTION              |  |   |                                      |
| #description          | catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  |   |                                      |
| #pathway              | glycogen/starch degradation  |   |                                      |
| CLASSIFICATION        | #superfamily alpha-amylase, amyloliquefaciens type; alpha-amylase core homology  |   |                                      |
| KEYWORDS              | extracellular protein; glycosidase; heat-stable protein; hydrolase; polysaccharide degradation   |   |                                      |
| FEATURE               |  |   |                                      |
| 1-34                  | #domain signal sequence  | #status predicted   | #label SIG\                          |
| 35-548                | #product alpha-amylase   | #status experimental                                      | #label MAT\                          |
| 235-368               | #domain alpha-amylase core homology  | #label AM\  |                                      |
| 139,237,272           | #binding_site calcium (Asp, Asp, His)  | #status predicted   |                                      |
| 268,298,365           | #active_site Asp, Glu, Asp   | #status predicted   |                                      |
| SUMMARY               | #length 548  | #molecular-weight 62585                                   | #checksum 3079                       |
| Query Match           | 65.2%  | Score 2524;   | DB 1; Length 548;                    |
| Best Local Similarity | 65.2%;   | Pred. No. 0.00e+00;                                       |                                      |
| Matches               | 330;   | Conservative  | 81; Mismatches 87; Indels 8; Gaps 8; |
| Db                    | 15   | LLAPLLTAS-LF-CPTGR-HAKAAPFNGTMMQYFEWYLPDDGTLTWKVANEANLSLG | 71                                   |
| QY                    | 11   | LTLLAVAVLFPTYTEPAQHNGT-NGTMMQYFEWHLPNDRHNRDLRDAANLKS      | 69                                   |
| Db                    | 72   | ITALSLPPAYKGTGRSDVGYVDLYDLGFENQGTGRTKYGTAQYIQ-AIQAAHAGM   | 130                                  |
| QY                    | 70   | ITAWYIPPAWKGTSNDVGYGAYDLYDLGFENQGTGRTKYGTRSQ-LOGAVTSLKNGI | 128                                  |

```
Db 185 NENVDDFEGNXYLNFANIDYNHPDVRREMDWGKWLIDTLQCGFRDLDAIKHINHEIFK 244
QY 219 DWEVDIENGNDYLYMAYADIMDHPVINELRNNGWVWYTNLNDGFRIDAVKHIKYSYTR 278
Db 245 EFAAEMIRKRGQDFYIVGEFWSNLDACREFFLDVVDYQIDFLDYSLYKHEASLKGDRDF 304
QY 279 DWLTHVRNTTGKPMFAEWFKNDLAAIENYLNKTSNNHSVDFDPLHYNLYNASNGGYF 338
Db 305 DLSKIFDDTLVQTHPTAVTFVNDHDSOPHEALESWIGDMFKSPAYALTLLRRDGYPVVF 364
QY 339 DMENILNGSVVQKPIHAVTFVNDHDSOPGEALESFVQSWFKPLAYALILTRQGGYPSVF 398
Db 365 YGDIYIGGPEPDGKKEIIDLARSARCAYGEQEDYFDHANTIGVWRGVEIEGSC 424
QY 399 YGDIYIGP-THGVPSMKSIDPLQARQYAYGTQHDYFDHDIIGTREGDSSHPNSGL 457
Db 425 AVVISNGDDGKRMFGEHRAGEVWDLTKSCDDQITIEDGWTAFHVCGGGVSW 480
QY 458 ATMSDGPCKNMYVKGKAGQWRDITGNRSCTVTINADGNGFTVNGGAVSW 513

RESULT 9
ENTRY 1BPLB #type complete
TITLE alpha-1,4-glucan-4-glucanohydrolase (EC 3.2.1.1), chain B -
ALTERNATE_NAMES alpha-1,4-glucan-4-glucanohydrolase
PDB_TITLE Bacillus licheniformis
ORGANISM alpha-amylase (bla)
#note glycosyltransferase
#formal_name Bacillus licheniformis
REFERENCE ATCC: 27811
#authors A65206
#submission Machius, M.; Wiegand, G.; Huber, R.
#cross-references PDB:1BPL
#journal J. Mol. Biol. (1995) 246:545-559
#book Crystal structure of calcium-depleted Bacillus licheniformis
#title alpha-amylase at 2.2 A resolution.
COMMENT Resolution: 2.2 angstroms
COMMENT Determination: X-ray diffraction
COMMENT R-value: no refinement
KEYWORDS 4-glucan-4-glucanohydrolase; alpha-1; alpha-amylase
FEATURE glycosyltransferase; glycosyltransferase
2-5 #region helix (right hand 3-10)
14-31 #region helix (right hand 3-10)
41-43 #region helix (right hand 3-10)
46-60 #region helix (right hand 3-10)
75-84 #region helix (right hand 3-10)
94-105 #region helix (right hand 3-10)
112-116 #region helix (right hand 3-10)
120-123 #region helix (right hand 3-10)
125-127 #region helix (right hand 3-10)
149-161 #region helix (right hand 3-10)
171-175 #region helix (right hand 3-10)
189-201 #region helix (right hand 3-10)
250-252 #region helix (right hand 3-10)
35-38,65-68 #region beta sheet
207-210,216-221, #region beta sheet
232-237,285-288 #region beta sheet
242-247,275-280 #region beta sheet
7-9 #region beta sheet
39,69,136 #length 290 #molecular-weight 33024 #checksum 5042
SUMMARY
Query Match 40.2%; Score 1558; DB 5; Length 290;
Best Local Similarity 67.5%; Pred. No. 1.00e-259;
Matches 195; Conservative 51; Mismatches 40; Indels 3; Gaps 2;
Db 1 YDLYMADIDYDHPDVAEIKRWGTWYANELQDGFRLDAVKHKFSLRDWNVHREKT 60
QY 229 YDLYMADIDMDHPVINELRNNGWVWYTNLNDGFRIDAVKHIKYSYTRDWTHTVRNT 288
```

```
Db 61 GKEMFTVAEYWDNDLGALENYLNKTNFNHSVDFDPLHYPFHAASTOGGDMRKLNSTV 120
QY 289 GKPMFAEAEFKNDLAAIENYLNKTSNNHSVDFDPLHYNLYNASNGGYDMRNLGVS 348
Db 121 VSKHPLKAVTFVNDHDSOPGEALESFVQSWFKPLAYALILTRQGGYPSVFYDYG 180
QY 349 VOKHPIHAVTFVNDHDSOPGEALESFVQSWFKPLAYALILTRQGGYPSVFYDYG 406
Db 181 SOREIPALKHKEPILKARKQYAGAOHDYFDHDIIGTREGDSSVANSGLAALITDGP 240
QY 407 TH-GVPSMKSIDPLQARQYAYGTQHDYFDHDIIGTREGDSSHPNSGLATIMSDGP 465
Db 241 GGAKRMYVGRONAGETHDITGNRSCTVTINADGNGFTVNGGAVSW 289
QY 456 GGNKMYVKGKAGQWRDITGNRSCTVTINADGNGFTVNGGAVSW 514

RESULT 10
ENTRY 1VJS2 #type fragment
TITLE alpha-amylase (EC 3.2.1.1), fragment 2 - Bacillus
ALTERNATE_NAMES bla
PDB_TITLE structure of alpha-amylase precursor
ORGANISM formal_name Bacillus licheniformis
#note strain bacillus licheniformis, ATCC: 27811
REFERENCE A66860
#authors Song, H.K.; Hwang, K.Y.; Chang, C.; Suh, S.W.
#submission submitted to the Brookhaven Protein Data Bank, October 1996
#cross-references PDB:1VJS
REFERENCE TN032246
#journal in Enzymes for Carbohydrate Engineering (In: Prog.
#book Biotechnol., V.12), Park, K.H.
#title Crystal structure of bacillus licheniformis alpha-amylase at
REFERENCE TN032247
#authors Lee, S.Y.; Kim, S.; Sweet, R.M.; Suh, S.W.
#journal Arch. Biochem. Biophys. (1991) 291:255
#title Crystallization and a preliminary X-ray crystallographic
study of alpha-amylase from bacillus licheniformis.
COMMENT Resolution: 1.7 angstroms
COMMENT Determination: X-ray diffraction
COMMENT R-value: 0.199
KEYWORDS carbohydrate metabolism; glycosidase; hydrolase
FEATURE
2-5 #region helix (right hand 3-10)
14-31 #region helix (right hand 3-10)
41-43 #region helix (right hand 3-10)
46-60 #region helix (right hand 3-10)
75-84 #region helix (right hand 3-10)
94-105 #region helix (right hand 3-10)
112-116 #region helix (right hand 3-10)
120-123 #region helix (right hand 3-10)
125-127 #region helix (right hand 3-10)
149-161 #region helix (right hand 3-10)
171-175 #region helix (right hand 3-10)
189-201 #region helix (right hand 3-10)
250-252 #region helix (right hand 3-10)
129-131,166-170, #region beta sheet
35-38,65-68 #region beta sheet
207-210,216-221, #region beta sheet
232-237,285-289 #region beta sheet
242-247,275-280 #region beta sheet
7-9 #region beta sheet
#length 290 #checksum 5042
SUMMARY
Query Match 40.2%; Score 1558; DB 5; Length 290;
Best Local Similarity 67.5%; Pred. No. 1.00e-259;
Matches 195; Conservative 51; Mismatches 40; Indels 3; Gaps 2;
Db 1 YDLYMADIDYDHPDVAEIKRWGTWYANELQDGFRLDAVKHKFSLRDWNVHREKT 60
QY 229 YDLYMADIDMDHPVINELRNNGWVWYTNLNDGFRIDAVKHIKYSYTRDWTHTVRNT 288
```

|                       |     |   |                                |
|-----------------------|-----|---|--------------------------------|
| Qy                    | 229 | YDYLMTADIMDHPVINELRNWCVVYTNILNDGFRIDAVKHKYSYTRDMLTHVRNIT          | 288                            |
| Db                    | 61  | GRKMFVAEYNDLIGALENYLNKTNFENSVEDVPLHVFQFHAOSTGGYDMRKLNSV           | 120                            |
| Qy                    | 289 | GRKMFVAEYNDLIGALENYLNKTNFENSVEDVPLHVFQFHAOSTGGYDMRKLNSV           | 120                            |
| Db                    | 121 | VSKHPLKAVTFVNDHDTQPOGLESYQVTFKPLAYAFILTRRESGYPOVFGYDYG            | 348                            |
| Qy                    | 349 | VOKHPHIAVTFVNDHDSQGEALSFVQSMFKPLAYALILTRREGYPSVFGYDYG             | 180                            |
| Db                    | 181 | SOREIPALKKIEPILKARKQYVAGQHDYFDHHDIVGTREGDSVANSGLAALITDGP          | 406                            |
| Qy                    | 407 | TH-CVPSMKSIDPLQARQYATGQHDYFDHHDIVGTREGDSVANSGLAALITDGP            | 240                            |
| Db                    | 241 | GGAKRWYGRQAGETHDITGNRSEPVWINSRGWGFHNGSGYSIYV                      | 465                            |
| Qy                    | 466 | GNKNWYVGHKAGQWRDITGNSGVTINADGNGFTVNGAVSVV                         | 289                            |
| RESULT                | 11  |   |                                |
| ENTRY                 |     | B45738  | #type complete                 |
| TITLE                 |     | alpha-amylase (EC 3.2.1.1), cytoplasmic - Salmonella typhimurium  |                                |
| ALTERNATE_NAMES       |     | 1,4-alpha-D-glucan glucanohydrolase                               |                                |
| ORGANISM              |     | #formal_name Salmonella typhimurium                               |                                |
| DATE                  |     | 07-Apr-1994   | #sequence_revision 18-Aug-1995 |
| ACCESSIONS            |     | 05-Sep-1997   | #text_change                   |
| REFERENCE             |     | B45738  |                                |
| #authors              |     | Raha, M.; Kawagishi, I.; Mueller, V.; Kihara, M.; Macnab, R.M.    |                                |
| #journal              |     | J. Bacteriol. (1992) 174:6644-6652                                |                                |
| #title                |     | Escherichia coli produces a cytoplasmic alpha-amylase, amyA.      |                                |
| #accession            |     | B45738  |                                |
| #molecule_type        |     | DNA   |                                |
| #residues             |     | 1-494   | #label RAH                     |
| GENETICS              |     | ##cross-references GB:L01643; NID:g154043; PID:g154045            |                                |
| #gene                 |     | amyA  |                                |
| FUNCTION              |     |   |                                |
| #description          |     | catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds |                                |
| #pathway              |     | glycogen/starch degradation                                       |                                |
| CLASSIFICATION        |     | #superfamily alpha-amylase, amylolytic                            |                                |
| KEYWORDS              |     | alpha-amylase core homology                                       |                                |
| FEATURE               |     | glycosidase; hydrolase; polysaccharide degradation                |                                |
| 202-335               |     | #domain alpha-amylase core homology                               |                                |
| 239,265,332           |     | #active_site His, Glu, Asp  |                                |
| SUMMARY               |     | #length 494 #molecular-weight 56496 #checksum 9082                |                                |
| Query Match           |     | 36.28; Score 1403; DB 1; Length 494;                              |                                |
| Best Local Similarity |     | 42.58; Pred. No. 2,028-230;                                       |                                |
| Matches               |     | 209; Conservative 106; Mismatches 159; Indels 18; Gaps 15;        |                                |
| Db                    | 3   | NFTLQVHHYYPDGGKLSLAERADGLINDIGINVWLPACKGASGYSVGYTDYDLF            | 62                             |
| Qy                    | 37  | NGTMQVFEWHLPDNGHNHNRDPAALKSIGITAVWIPAKGTSQ-N-DVGIGAYDLY           | 95                             |
| Db                    | 63  | DLGEFDQKGTATYKCGKQLLTAIDALKKNIAVLDDVYVNHKGADEKERIRVQV             | 122                            |
| Qy                    | 96  | DLGEFNKGTVTKGTGTSQLOGAVTSLKNGIGQYGVYVNHKGADEKERIRVQV              | 122                            |
| Db                    | 123 | DDRTQDDNIIIEGWTYTPPARAGYSNFW-D-YHCFSGIDHIENPD-EDGIFKTVND          | 180                            |
| Qy                    | 156 | SNRQSEISYEITAWTKDFPGRGNTHSNFKRWYH-FDGTDMDSQRLQNKIYF-RG            | 213                            |
| Db                    | 181 | YTGDSGNDQVDEMGENDYLMGENDFRHATEIKYWARVMEQTHCDGFRDLAVKHI            | 240                            |
| Qy                    | 214 | -TGKAWDEVDIENGNDYLMYADIMDRIPEINELRNWGYTNTLNDGFRDLAVKHI            | 272                            |
| Db                    | 241 | PAWFKIEWIHQVAPKPLFYAEVSHVEDVKLTQIDQVQDKTMLFDAPLQMKFEAS            | 300                            |

QY 37 NGLMAYFEWHPNDGQHWKRLQNDSDAYLAHGIITAVWIPPAWKGTSON-DVGAYDLY 95  
DB 63 DLGEFDKQSIPTKYGDKAQLAALDAIKRNDIAVLDDVVVNNHKGADKEAIRVQRYNA 122  
QY 96 DLGEFNKGKTVRTKYGRSLOQAVTSLKNNGIQVYGVVNNHKGADGTEMWNAVEVNR 155  
DB 123 DRTQ-IDELIEEGWTRTFPARAQYQFIWDFKCFSGIDHENDP-EDGIFKIVND 180  
QY 156 SNRQETSGEYTI-EAWTKFDFPGRGNTSHNFKRWYHFDGTDQDQSLQNLKIKF-RG 213  
DB 181 YTEGNDQVDDDELNGNDYILNGENIDRNHVAETEEIKYARWVMEQOTCDGFRLDVAKHI 240  
QY 214 -TKANDWEVDIENGNTDYILMYADIDMDHPEVINELNRNGWVYNTNLNGFRIDAVAKHI 272  
DB 241 PAWFYKEWIEHVOVAPKPLFIYAEYWSHEVDKLTQYIDQVEGKTMFLDAPLQMKFEAS 300  
QY 273 KYSTRDMLTHVRTTKPKPFAVAEFKNDLAAENLNKTSWNHSDVDFPLHNLINYNAS 332  
DB 301 RMGRDYDMTOIFGTLVADPFHATVILVANDHOTOPLQALEAPVPWPKPLAYALILREN 360  
QY 333 NSGGEYFDMRNLNGSVYQKHPHIAVTFVNDHDSQGEALESFVQSWFKPLAYALILREQ 392  
DB 361 GVPSVEPDLYGAHYEDVGGQYIPIDMPIEOLDELILARQFAHGVQLFDFHPNCI 420  
QY 393 GYPSVFYGD-Y---Y-GIPTHG-V-P-SMK--SKIDPLQARQYAYGTQHDYDFHDHII 442  
DB 421 AFSRSGDFEPFGC-VV-VMSNGDDGKETHILGENYGNKTRDFLGNROERVVTIDENGSEAT 478  
QY 443 GWTREGDSSHPNSGLATIMSDGPGCNKMYGKHKAGQVWDITGNRSCTVINADGWN 502  
DB 479 FFCNGGSVSVWV 490  
QY 503 FTVNGGAVSVWV 514

RESULT 13  
ENTRY TITLE  
ALTERNATE\_NAMES 1BPLA #type complete  
PDB\_TITLE alpha-1,4-glucan-4-glucanohydrolase (EC 3.2.1.1), chain A -  
ORGANISM Bacillus licheniformis  
#note glycosyltransferase  
#authors #formal\_name Bacillus licheniformis  
#cross-references ATCC: 27811  
#submission Machius, M.; Wiegand, G.; Huber, R.  
#cross-references PDB:1BPL submitted to the Brookhaven Protein Data Bank, July 1995  
#authors Machius, M.; Wiegand, G.; Huber, R.  
#cross-references PDB:1BPL J. Mol. Biol. (1995) 246:545-559  
#title Crystal structure of calcium-depleted Bacillus licheniformis  
alpha-amylase at 2.2 A resolution.  
COMMENT Resolution: 2.2 angstroms  
COMMENT Determination: X-ray diffraction  
COMMENT R-value: no refinement  
KEYWORDS 4-glucan-4-glucanohydrolase; alpha-1; alpha-amylase  
glycosyltransferase; glycosyltransferase  
FEATURE  
19-25 #region helix (right hand alpha)\  
27-32 #region helix (right hand alpha)\  
78-91 #region helix (right hand alpha)\  
155-157 #region helix (right hand alpha)\  
5-7,37-39,94-99 #region beta sheet\  
131-135 111-117, #region beta sheet\  
173-175 #region beta sheet\  
158-161 #region beta sheet\  
SUMMARY #length 179 #molecular-weight 20547 #checksum 2012

Query Match 26.0%; Score 1006; DB 5; Length 179;  
Best Local Similarity 70.9%; Pred. No. 6.06e-156;  
Matches 127; Conservative 26; Mismatches 25; Indels 1; Gaps 1;

DB 2 NGLMAYFEWHPNDGQHWKRLQNDSDAYLAHGIITAVWIPPAWKGTSONDVGAYDLYD 61  
QY 37 NGLMAYFEWHPNDGQHWKRLQNDSDAYLAHGIITAVWIPPAWKGTSONDVGAYDLYD 96  
DB 62 LGEPHOKGTVRTKYTGKGLQSAIKSLHSRDLNVYGVVNNHKGADGTEMWNAVEVNR 121  
QY 97 LGEPHOKGTVRTKYTGKGLQSAIKSLHSRDLNVYGVVNNHKGADGTEMWNAVEVNR 156  
DB 122 DNRRVVISGEHLIKANTHFFPGRGSTYSDFKWHWYHFDGTDWDESRLK-NRIYKFOGKA 179  
QY 157 NRNQETSGEYTI-EAWTKFDFPGRGNTSHNFKRWYHFDGTDQDQSLQNLKIKF-RGTG 215

RESULT 14  
ENTRY TITLE  
ALTERNATE\_NAMES 1VJS1 #type fragment  
PDB\_TITLE alpha-amylase (EC 3.2.1.1), fragment 1 - Bacillus  
ORGANISM licheniformis  
#note structure of alpha-amylase precursor  
#authors #formal\_name Bacillus licheniformis  
#cross-references strain bacillus licheniformis, ATCC: 27811  
#submission A6860  
#cross-references Song, H.K.; Hwang, K.Y.; Chang, C.; Suh, S.W.  
#cross-references submitted to the Brookhaven Protein Data Bank, October 1996  
#cross-references PDB:1VJS  
#cross-references TN032244  
#authors Song, H.K.; Hwang, K.Y.; Chang, C.; Suh, S.W.  
#book in Enzymes for Carbohydrate Engineering (In: Prog. Biotechnol., V.12), Park, K.H.  
#title Crystal structure of bacillus licheniformis alpha-amylase at 1.7 a resolution.  
REFERENCE TN032245  
#authors Lee, S.Y.; Kim, S.; Sweet, R.M.; Suh, S.W.  
#journal Arch. Biochem. Biophys. (1991) 291:255  
#title Crystallization and a preliminary X-ray crystallographic study of alpha-amylase from bacillus licheniformis.  
COMMENT Resolution: 1.7 angstroms  
COMMENT Determination: X-ray diffraction  
COMMENT R-value: 0.199  
KEYWORDS carbohydrate metabolism; glycosidase; hydrolase  
FEATURE  
19-25 #region helix (right hand alpha)\  
27-32 #region helix (right hand alpha)\  
78-91 #region helix (right hand alpha)\  
155-157 #region helix (right hand alpha)\  
5-7,37-39,94-99 #region beta sheet\  
173-175,109-117, #region beta sheet\  
131-138 #region beta sheet\  
158-161 #region beta sheet\  
SUMMARY #length 179 #checksum 2012

Query Match 26.0%; Score 1006; DB 5; Length 179;  
Best Local Similarity 70.9%; Pred. No. 6.06e-156;  
Matches 127; Conservative 26; Mismatches 25; Indels 1; Gaps 1;

DB 2 NGLMAYFEWHPNDGQHWKRLQNDSDAYLAHGIITAVWIPPAWKGTSONDVGAYDLYD 61  
QY 37 NGLMAYFEWHPNDGQHWKRLQNDSDAYLAHGIITAVWIPPAWKGTSONDVGAYDLYD 96  
DB 62 LGEPHOKGTVRTKYTGKGLQSAIKSLHSRDLNVYGVVNNHKGADGTEMWNAVEVNR 121  
QY 97 LGEPHOKGTVRTKYTGKGLQSAIKSLHSRDLNVYGVVNNHKGADGTEMWNAVEVNR 156  
DB 122 DNRRVVISGEHLIKANTHFFPGRGSTYSDFKWHWYHFDGTDWDESRLK-NRIYKFOGKA 179  
QY 157 NRNQETSGEYTI-EAWTKFDFPGRGNTSHNFKRWYHFDGTDQDQSLQNLKIKF-RGTG 215

RESULT 15  
ENTRY TITLE  
ALTERNATE\_NAMES 1VJS1 #type fragment  
PDB\_TITLE alpha-amylase (EC 3.2.1.1), fragment 1 - Bacillus  
ORGANISM licheniformis  
#note structure of alpha-amylase precursor  
#authors #formal\_name Bacillus licheniformis  
#cross-references strain bacillus licheniformis, ATCC: 27811  
#submission A6860  
#cross-references Song, H.K.; Hwang, K.Y.; Chang, C.; Suh, S.W.  
#cross-references submitted to the Brookhaven Protein Data Bank, October 1996  
#cross-references PDB:1VJS  
#cross-references TN032244  
#authors Song, H.K.; Hwang, K.Y.; Chang, C.; Suh, S.W.  
#book in Enzymes for Carbohydrate Engineering (In: Prog. Biotechnol., V.12), Park, K.H.  
#title Crystal structure of bacillus licheniformis alpha-amylase at 1.7 a resolution.  
REFERENCE TN032245  
#authors Lee, S.Y.; Kim, S.; Sweet, R.M.; Suh, S.W.  
#journal Arch. Biochem. Biophys. (1991) 291:255  
#title Crystallization and a preliminary X-ray crystallographic study of alpha-amylase from bacillus licheniformis.  
COMMENT Resolution: 1.7 angstroms  
COMMENT Determination: X-ray diffraction  
COMMENT R-value: 0.199  
KEYWORDS carbohydrate metabolism; glycosidase; hydrolase  
FEATURE  
19-25 #region helix (right hand alpha)\  
27-32 #region helix (right hand alpha)\  
78-91 #region helix (right hand alpha)\  
155-157 #region helix (right hand alpha)\  
5-7,37-39,94-99 #region beta sheet\  
173-175,109-117, #region beta sheet\  
131-138 #region beta sheet\  
158-161 #region beta sheet\  
SUMMARY #length 179 #checksum 2012

Query Match 26.0%; Score 1006; DB 5; Length 179;  
Best Local Similarity 70.9%; Pred. No. 6.06e-156;  
Matches 127; Conservative 26; Mismatches 25; Indels 1; Gaps 1;



```

DATE          05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change
ACCESSIONS    31-Oct-1997
REFERENCE      A91760; A91759; A19506
#authors      Sachdev, O.; Friedberg, F.
#journal      Int. J. Pept. Protein Res. (1981) 18:228-236
#cross-references MUID:82189140
#accession    A91760
#molecule_type protein
#residues     1-56;57-144 #label SAC
##note       residues 1-56 correspond to residues 229-287 of the
              complete sequence; residues 57-144 (without residues
              140-144) correspond to residues 335-397 of the
              complete sequence

REFERENCE      A91759
#authors      Detera, S.D.; Friedberg, F.
#journal      Int. J. Pept. Protein Res. (1981) 17:93-106
#cross-references MUID:81191186
#accession    A91759
#molecule_type protein
#residues     145-217 #label DET
##note       this sequence corresponds to residues 398-469 of the
              complete sequence

COMMENT       See ALBSN.
FUNCTION      catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic
              bonds
#pathway      glycogen/starch degradation
CLASSIFICATION #superfamily alpha-amylase, amyloliquefaciens type;
KEYWORDS      alpha-amylase core homology
              glycosidase; hydrolase; polysaccharide degradation
SUMMARY       #length 217 #checksum 7645

Query Match      8.9%; Score 344; DB 2; Length 217;
Best Local Similarity 60.8%; Pred. No. 1.87e-36;
Matches          59; Conservative 14; Mismatches 13; Indels 11; Gaps 7;

Db 118 AYWAFILTRREGYGVYGVGVGYMGYGTGTPKPEIPSLKDNIEPILKARKEYAYGPO 177
QY 383 AY-ALLITREQ-GYPSVYFGD---Y-YGI-PT--HGVPFMSKSIDPLQLQARTYAYGTQ 432
Db 178 HDYIDPH-VIGWTRGDSSAAKSLAALISDGPGGK 213
QY 433 HDYFDHHDIIIGWTRGDSSHPNSGLATIMSDGPGNK 469

Search completed: Tue Sep 15 14:17:58 1998
Job time : 81 secs.

```

\*\*\*\*\*  
M P S R L H  
\*\*\*\*\* (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Sep 15 14:13:48 1998; MasPar time 22.56 Seconds  
Tabular output not generated. 370.088 Million cell updates/sec

Title: >US-08-952-741-2  
Description: (1-516) from US08952741.pep  
Perfect Score: 3873  
Sequence: 1 MKLHRIISVLLTLLAVAV.....ADGWNFTYNGGAVSVWVKQ 516

Scoring table: PAM 150  
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq32  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29

Statistics: Mean 35.739; Variance 162.847; scale 0.219

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description                  | Pred. No. |
|------------|-------|-------------|--------|----|------------------------------|-----------|
| 1          | 3873  | 100.0       | 516    | 24 | W11326 Alkaline liquefying a | 0.00e+00  |
| 2          | 3539  | 91.4        | 485    | 27 | W31499 Bacillus sp. alpha am | 0.00e+00  |
| 3          | 3537  | 91.3        | 485    | 20 | W12131 Alpha-amylase variant | 0.00e+00  |
| 4          | 3536  | 91.3        | 485    | 20 | W12120 Alpha-amylase variant | 0.00e+00  |
| 5          | 3536  | 91.3        | 485    | 20 | W12125 Alpha-amylase variant | 0.00e+00  |
| 6          | 3536  | 91.3        | 485    | 20 | W12128 Alpha-amylase variant | 0.00e+00  |
| 7          | 3536  | 91.3        | 485    | 20 | W12118 Alpha-amylase variant | 0.00e+00  |
| 8          | 3536  | 91.3        | 485    | 20 | W12122 Alpha-amylase variant | 0.00e+00  |
| 9          | 3536  | 91.3        | 485    | 20 | W12123 Alpha-amylase variant | 0.00e+00  |
| 10         | 3536  | 91.3        | 485    | 20 | W12130 Alpha-amylase variant | 0.00e+00  |
| 11         | 3536  | 91.3        | 485    | 20 | W12129 Alpha-amylase variant | 0.00e+00  |
| 12         | 3536  | 91.3        | 485    | 20 | W12127 Alpha-amylase variant | 0.00e+00  |
| 13         | 3536  | 91.3        | 485    | 20 | W12126 Alpha-amylase variant | 0.00e+00  |
| 14         | 3536  | 91.3        | 485    | 20 | W12124 Alpha-amylase variant | 0.00e+00  |
| 15         | 3536  | 91.3        | 485    | 20 | W12108 Alpha-amylase variant | 0.00e+00  |
| 16         | 3536  | 91.3        | 485    | 20 | W12119 Alpha-amylase variant | 0.00e+00  |
| 17         | 3536  | 91.3        | 485    | 20 | W12121 Alpha-amylase variant | 0.00e+00  |
| 18         | 3534  | 91.2        | 485    | 20 | W12114 Alpha-amylase variant | 0.00e+00  |

|    |      |      |     |    |                              |          |
|----|------|------|-----|----|------------------------------|----------|
| 19 | 3534 | 91.2 | 485 | 20 | W12117 Alpha-amylase variant | 0.00e+00 |
| 20 | 3533 | 91.2 | 485 | 20 | W12113 Alpha-amylase variant | 0.00e+00 |
| 21 | 3532 | 91.2 | 485 | 20 | W12105 Alpha-amylase variant | 0.00e+00 |
| 22 | 3532 | 91.2 | 485 | 20 | W12099 Alpha-amylase variant | 0.00e+00 |
| 23 | 3532 | 91.2 | 485 | 20 | W12106 Alpha-amylase variant | 0.00e+00 |
| 24 | 3532 | 91.2 | 485 | 20 | W12100 Alpha-amylase variant | 0.00e+00 |
| 25 | 3532 | 91.2 | 485 | 20 | W12098 Alpha-amylase variant | 0.00e+00 |
| 26 | 3532 | 91.2 | 485 | 20 | W12104 Alpha-amylase variant | 0.00e+00 |
| 27 | 3532 | 91.2 | 485 | 20 | W12102 Alpha-amylase variant | 0.00e+00 |
| 28 | 3532 | 91.2 | 485 | 20 | W12107 Alpha-amylase variant | 0.00e+00 |
| 29 | 3532 | 91.2 | 485 | 20 | W12103 Alpha-amylase variant | 0.00e+00 |
| 30 | 3529 | 91.1 | 485 | 20 | W12110 Alpha-amylase variant | 0.00e+00 |
| 31 | 3529 | 91.1 | 485 | 20 | W12109 Alpha-amylase variant | 0.00e+00 |
| 32 | 3528 | 91.1 | 485 | 20 | W12111 Alpha-amylase variant | 0.00e+00 |
| 33 | 3528 | 91.1 | 485 | 20 | W12144 Alpha-amylase variant | 0.00e+00 |
| 34 | 3528 | 91.1 | 485 | 20 | W12112 Alpha-amylase variant | 0.00e+00 |
| 35 | 3528 | 91.1 | 485 | 20 | W12101 Alpha-amylase variant | 0.00e+00 |
| 36 | 3513 | 90.7 | 485 | 16 | R81835 Bacillus sp. alkaline | 0.00e+00 |
| 37 | 3505 | 90.5 | 483 | 20 | W12116 Alpha-amylase variant | 0.00e+00 |
| 38 | 3503 | 90.4 | 483 | 20 | W12142 Alpha-amylase variant | 0.00e+00 |
| 39 | 3503 | 90.4 | 483 | 20 | W12141 Alpha-amylase variant | 0.00e+00 |
| 40 | 3501 | 90.4 | 483 | 20 | W12115 Alpha-amylase variant | 0.00e+00 |
| 41 | 3503 | 90.4 | 485 | 20 | W12132 Alpha-amylase variant | 0.00e+00 |
| 42 | 3495 | 90.2 | 483 | 20 | W12133 Alpha-amylase variant | 0.00e+00 |
| 43 | 3489 | 89.6 | 483 | 20 | W12134 Alpha-amylase variant | 0.00e+00 |
| 44 | 3464 | 89.4 | 483 | 20 | W12135 Alpha-amylase variant | 0.00e+00 |
| 45 | 3453 | 89.2 | 483 | 20 | W12136 Alpha-amylase variant | 0.00e+00 |

ALIGNMENTS

RESULT 1

ID W11326 standard; Protein; 516 AA.  
AC W11326:  
DT 11-NOV-1997 (first entry)  
DE Alkaline liquefying alpha-amylase.  
KW Alkaline liquefying alpha-amylase; Bacillus; alpha-amylase; detergent;  
KW starch-related polysaccharide; hydrolysis; enzyme; surfactant resistance;  
KW alpha-1,4-glucosidic link; alkalophilic Bacillus; laundry detergent;  
OS dish-washing detergent; starch.  
OS Bacillus species KSM-AP1378.  
PN WO9700324-A1.  
PD 03-JAN-1997.  
PF 14-JUN-1996; J01641.  
PR 14-JUN-1995; JP-147257.  
PA (KAOS) KAO CORP.  
PI Ara K, Hatada Y, Ito S, Kawai S, Ozaki K;  
DI WPI; 97-118708/11.  
DR N-PSDB; T51339.  
PT DNA encoding alkaline liquefying alpha-amylase - useful in  
PT dish-washing and laundry detergents for removal of starch dirt  
PS Claim 2: Page 23-26; 40pp; English.  
CC This sequence represents an alkaline liquefying alpha-amylase.  
CC Alpha-amylase is an enzyme that acts on starch-related polysaccharides,  
CC hydrolysing the alpha-1,4-glucoside bond of the polysaccharide molecule.  
CC Alkaline liquefying alpha-amylases exhibit resistance to surfactants used  
CC in detergents, and decompose starch or starch-related polysaccharides in  
CC a highly random manner. The Bacillus species KSM-AP1378, from which this  
CC sequence was isolated, is an alkalophilic Bacillus strain. It was  
CC isolated from soil in the vicinity of the city of Tochigi. The enzyme is  
CC useful in improving the efficiency of dish-washing and laundry  
CC detergents, particularly on starch dirt.  
SQ Sequence 516 AA;

Query Match 100.0%; Score 3873; DB 24; Length 516;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mklhriisvlltlllavavlfpytmepaqahhngtngtmqyfewhlpdngnwnrlrd 60

QY 1 MKLHRIISVLLTLLAVAVLFPPYTMPEAQAHNGTNGTMMQYFEWHLPDNGNHNRLRD 60

Db 61 daanlkskgitavwipawktsqndvgaydyldlgefnkgvtvtrkysqlgav 120

```
QY 61 DAANLKSIGITAVWIPPAWKTQSDYGYGAYDLYDLGEFNQKGTVRTKYGTRSQLQAV 120
Db 121 tslnngiqvvgdvvmhkhgagadtemnavvevnsrnrnqeisgeytieawtkfdpgrg 180
QY 121 TSLKNNGIOVGVVMMHKGAGADEMNAVVEVNSRNRNOEISGEYTIETAWTKFDFPGR 180
Db 181 nthsfnkwrwhfdgtgdwdsrqlqkikyfrgtgkdwedvdiengnydylmvdadmd 240
QY 181 NTHSNFKWRWHFDGTGDWDSRQLQKIKYFRGTGKADWEVDIENGNYDYLMDADMD 240
Db 241 hpevineelnrgvwyntlnldgfridavkhikysytrdwlthvtrnttggkpmfavaefwk 300
QY 241 HPEVINEELNRGWVYNTLNLDGFRIDAVKHIKYSYTRDWLTHVTRNTTGGKPMFAVAEFWK 300
Db 301 ndlaaenylnktswnhsvfdvplhynlnasnggyfdmrnlngsvvqkhpiahvtfv 360
QY 301 NDLAATENYLNKTSWNHSHVFDVPLHYNLYNASNSGGYFDMRNILNGSVVQKHPIAHVTFV 360
Db 361 dnhdsqgealesfvgsfwkplayalltregqpsvfygdygipthgvpmskksidpl 420
QY 361 DNHDSQGEALESFVGSFWKPLAYALLTREGQPSVFGYDYGIPTHGVPMSKSKIDPL 420
Db 421 lqarqtyaygtqhdvfdhddiigtregdshpnsnglatimsdpgpgnkmyvgkhhkag 480
QY 421 LQARQTYAYGTQHDYFDHDDIIGTREGDSSHPNSGLATIMSDPGPGNKMYVGKHKAG 480
Db 481 vwrldtgnrsgtgtvtnadgwnftvnggavsvvkwk 516
QY 481 VWRDITGNRSGTGTINADGWNFTVNGGAVSVVWKQ 516
```

## RESULT 2

```
ID W31499 standard; protein; 485 AA.
AC W31499;
DT 08-APR-1998 (first entry)
DE Bacillus sp. alpha amylase.
KW Alpha amylase; hard surface cleaning; dishwashing; laundry.
OS Bacillus sp.
PN W09732961-A2.
PD 12-SEP-1997.
PF 04-MAR-1997; U03635.
PR 07-MAR-1996; WO-U03276.
PA (PROC ) PROCTER & GAMBLE CO.
PI Baek AC, Jones IA, Ohtani R, Pramod K, Rai S,
PI Showell MS, Ward G;
PI WPI; 97-457524/42.
PT Detergent compositions for hard surface cleaning and laundry use -
PT contains Bacillus derived alpha amylase with improved
PT thermostability, reduced calcium ion dependency etc.
PS Claim 1; Pages 86-87; 97pp; English.
CC The present sequence is a Bacillus sp. alpha amylase with a
CC specific activity at least 25% higher than that of Termamyl (RTM)
CC at 25 to 55 degrees C at pH 8 to 10, measured by the Phadebas (RTM)
CC activity test. It is of use in hard surface cleaning, hand or
CC machine dishwashing and laundry at a temperature of 10 to 25
CC degrees C. Improved cleaning, stain removal and fabric care are
CC obtained by using it at a concentration of 0.00018 to 0.06%
SQ Sequence 485 AA;
```

```
Query Match 91.4%; Score 3539; DB 27; Length 485;
Best Local Similarity 95.1%; Pred. No. 0.00e+00;
Matches 461; Conservative 18; Mismatches 6; Indels 0; Gaps 0;
```

```
Db 1 hngtngtmmygfewylpndgnhwnrirdaenlkskgitavppawkgtsqndvgga 60
QY 32 HHNTGTMTMYGFEMHLPNDGNHWNRRDDAANLKSIGITAVIPPAWKGTSQNDVGGA 91
Db 61 ydlldlgefngkgtvrtkygtrnglqaavtslknglqvgydvvmhkhgagdeivnav 120
QY 92 YDLIDLGEFNGKGTVRTKYGTRNGLQAAVTSLKNGLTQVGYDVVMHKGAGDEIVNAV 151
Db 121 evnrsnrnqetsgeysaeawtkfdpgrgnhssfkwrwhfdgtgdwdsrqlqkikyfk 180
```

```
QY 152 EVNSRNRNQEISGEYTIETAWTKFDFPGRGNTHSNFKWRWHFDGTGDWDSRQLQNKIKYFK 211
Db 181 rgtgkdwedvtenpnydylmvdadmdhpevihelnrgvwyntlnldgfridavkh 240
QY 212 RGTGKADWEVDIENGNYDYLMDADMDHPEVINEURNWGWYNTLNLDGFRIDAVKH 271
Db 241 ikysyfrdwlthvtrnttggkpmfavaefwkndlgaienylnktswnhsvfdvplhynlna 300
QY 272 IKYSYTRDWLTHVTRNTTGGKPMFAVAEFWKNDLGAIAENYLNKTSWNHSHVFDVPLHYNLYNA 331
Db 301 snsggyfdmrnlngsvvqkhpthavtfvndhdsqgealesfvqgfkplayalvtre 360
QY 332 SNSGGYFDMRNILNGSVVQKHPHIAVTFVNDHDSQGEALESFVQSFVKPLAYALILTRE 391
Db 361 qgypsvfygdygipthgvpmskksidpllqarqtyaygtqhdvfdhddiigtregdss 420
QY 392 QGYPSVFGYDYGIPTHGVPMSKSKIDPLLQARQTYAYGTQHDYFDHDDIIGTREGDSS 451
Db 421 hpnsnglatimsdpgpgnkmyvgknhkagqvrdrdtgnrtgtvtnadgwnfsvnggsvs 480
QY 452 HPNSGLATIMSDPGPGNKMYVGKHKAGQVWRDITGNRSGTGTINADGWNFTVNGGAVS 511
Db 481 vvwkq 485
QY 512 VWVKQ 516
```

## RESULT 3

```
ID W12131 standard; protein; 485 AA.
AC W12131;
DT 08-APR-1997 (first entry)
DE Alpha-amylase variant N106D.
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;
KW calcium ion dependency; alpha-amylolytic activity; washing composition;
KW textile desizing; papermaking; beer-making; ethanol production;
KW sweetener.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_difference 106 /label= N106D
FT WO9623873-A1.
PN 08-AUG-1996.
PF 05-FEB-1996; DK0056.
PR 03-FEB-1995; DK-000126.
PR 29-MAR-1995; DK-000336.
PR 29-SEP-1995; DK-001097.
PR 06-OCT-1995; DK-001121.
PA (NOVO ) NOVO-NORDISK AS.
PI Bisgard-frantzen H, Borchert T, Svendsen A;
PI WPI; 96-371423/37.
DR Alpha-amylase variants - with improved thermal and oxidation
PT stability and reduced calcium ion dependency
PS Claim 11; ; ilpp; English.
CC W12098-W12144 represent alpha-amylase variants of the invention. The
CC variants of the invention were created using site directed, or random,
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,
CC W12142 and W12144 are specifically variants of the alkaphilic bacillus
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have
CC improved thermal stability (such as at temperatures in the range of 40-70
CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion
CC dependency. The variants can also have increased alpha-amylolytic
CC activity (especially at pH values in the range of 8.5-10.5), and improved
CC binding of a particular substrate. These variant alpha-amylases also
CC possess improved specificity to a particular substrate, and/or improved
CC specificity with respect to cleavage of substrate. These sequences can
CC be used in detergent and washing compositions, and for textile desizing.
CC The alpha-amylase variants can also be used in papermaking and
CC beer-making processes. These variants can also be used in the production
CC of sweeteners and ethanol from starch.
SQ Sequence 485 AA;
```

Query Match 91.3%; Score 3537; DB 20; Length 485;  
Best Local Similarity 94.8%; Pred. NO. 0.00e+00;  
Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hhngtngtmmyfeywlpndgnhwnrlrddaanlkskgtawwippawkgtsqndvgyga 60  
QY 32 HHNGTNGTMQYFEWHLNDGNHWNRLRDDAANLKSIGTAVWIPPAWKGTSQNDVGYGA 91  
Db 61 ydlvdlgefngkgtvrtkygtrnqlqaavtslknngiqvygvdvnmhkggadgteivnav 120  
QY 92 YDLVDLGEFNGKGTVRTKYGTRSQLQGAVTSLKNNGIQVYGDVVMNHKGGADGTEMVNAV 151  
Db 121 evnsrnrnqetsgeyaieawtkfdpgrgnnhssfkwrwyhfdgtdwdgsrqlqkniyxf 180  
QY 152 EVNSRNRNQEISGEYIEAWTKFDPGRGNTHSNFKWRWYHFDGTDWDQSRQLQNKIYKF 211  
Db 181 rgtgkdwewdvengnydylymadvdmhpevnhelrnwgvwyntnlndgfridavkh 240  
QY 212 RGTGKADWEVDIENGNDYLYMADIDMDHPEVNHENLWGWVYNTNLNDGFRIDAVKH 271  
Db 241 ikysftrdwlthvrnttqkpmfavaefwkndlgaienylnktswnhsvfdvplhynlyna 300  
QY 272 IKYSTRDWLTHVRNTTQKPMFAVAEFWKNDLAAIENYLNKTSWNHSVFDVPLHYLYNA 331  
Db 301 snsggyydmrnlngsvvqkphthavtfvndhdsqgealesfvqgwkfplayalvtre 360  
QY 332 SNSGGYFDMRNLNGSVVQKPHIHAVTFVNDHDSQGEALESFVQSWFKPLAYALITRE 391  
Db 361 qgypsvfygdygipthgvpamkskidpqlqarqfaygtqhdhdiigwtregss 420  
QY 392 QGYPSVFGDYIGIPTHGVPMSKSIDPQLQARQIAYGTQHDYFDHDDIIGWTREGSS 451  
Db 421 hpnsclatimsdpggnkwmvvgkkgagvwrtditgnrtgtvtinadgwnfsvnggsvs 480  
QY 452 HPNSGLATIMSDPGGNKWMVVGKKGAGQVWRDITGNRSSTVTINADGWNFTVNGGAVS 511  
Db 481 vvwkq 485  
QY 512 VVWKQ 516  
RESULT 4  
ID W12120 standard; protein; 485 AA.  
AC W12120;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant K239R.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_difference 239 /label= K239R  
FT W09623873-A1.  
FN 08-AUG-1996.  
PD 05-FEB-1995; DK0056.  
PF 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO) NOVO-NORDISK AS.  
PI Bisgaard-frantzen H, Borchert T, Svendsen A;  
DR WPI: 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Claim 11; : ilpp: English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have

CC improved thermal stability (such as at temperatures in the range of 40-70  
CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
Best Local Similarity 94.8%; Pred. NO. 0.00e+00;  
Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hhngtngtmmyfeywlpndgnhwnrlrddaanlkskgtawwippawkgtsqndvgyga 60  
QY 32 HHNGTNGTMQYFEWHLNDGNHWNRLRDDAANLKSIGTAVWIPPAWKGTSQNDVGYGA 91  
Db 61 ydlvdlgefngkgtvrtkygtrnqlqaavtslknngiqvygvdvnmhkggadgteivnav 120  
QY 92 YDLVDLGEFNGKGTVRTKYGTRSQLQGAVTSLKNNGIQVYGDVVMNHKGGADGTEMVNAV 151  
Db 121 evnsrnrnqetsgeyaieawtkfdpgrgnnhssfkwrwyhfdgtdwdgsrqlqkniyxf 180  
QY 152 EVNSRNRNQEISGEYIEAWTKFDPGRGNTHSNFKWRWYHFDGTDWDQSRQLQNKIYKF 211  
Db 181 rgtgkdwewdvengnydylymadvdmhpevnhelrnwgvwyntnlndgfridavrh 240  
QY 212 RGTGKADWEVDIENGNDYLYMADIDMDHPEVNHENLWGWVYNTNLNDGFRIDAVKH 271  
Db 241 ikysftrdwlthvrnttqkpmfavaefwkndlgaienylnktswnhsvfdvplhynlyna 300  
QY 272 IKYSTRDWLTHVRNTTQKPMFAVAEFWKNDLAAIENYLNKTSWNHSVFDVPLHYLYNA 331  
Db 301 snsggyydmrnlngsvvqkphthavtfvndhdsqgealesfvqgwkfplayalvtre 360  
QY 332 SNSGGYFDMRNLNGSVVQKPHIHAVTFVNDHDSQGEALESFVQSWFKPLAYALITRE 391  
Db 361 qgypsvfygdygipthgvpamkskidpqlqarqfaygtqhdhdiigwtregss 420  
QY 392 QGYPSVFGDYIGIPTHGVPMSKSIDPQLQARQIAYGTQHDYFDHDDIIGWTREGSS 451  
Db 421 hpnsclatimsdpggnkwmvvgkkgagvwrtditgnrtgtvtinadgwnfsvnggsvs 480  
QY 452 HPNSGLATIMSDPGGNKWMVVGKKGAGQVWRDITGNRSSTVTINADGWNFTVNGGAVS 511  
Db 481 vvwkq 485  
QY 512 VVWKQ 516  
RESULT 5  
ID W12125 standard; protein; 485 AA.  
AC W12125;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant D199N.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_difference 199 /label= D199N  
FT W09623873-A1.  
FN 08-AUG-1996.  
PD 05-FEB-1995; DK0056.  
PF 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.

PR 06-OCT-1995; DK-001121.  
 PA (NOVO) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 stability and reduced calcium ion dependency  
 PS Claim 11: 111pp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 variants of the invention were created using site directed, or random,  
 mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 W12142 and W12144 are specifically variants of the alkalophilic bacillus  
 strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 improved thermal stability (such as at temperatures in the range of 40-70  
 degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 dependency. The variants can also have increased alpha-amylolytic  
 activity (especially at pH values in the range of 8.5-10.5), and improved  
 binding of a particular substrate. These variant alpha-amylases also  
 possess improved specificity to a particular substrate, and/or improved  
 specificity with respect to cleavage of substrate. These sequences can  
 be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 beer-making processes. These variants can also be used in the production  
 of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hngtngtmmqyfeywlpndgnhwnrlrddaanlkskgitavwippawkgtsqndvgyga 60  
 QY 32 HHNGTNGTMMQYFEWHLPLNDGNHNRLRDDAANLKSIGITAVWIPPAAWKGTSQNDVGYGA 91  
 Db 61 ydlldlgefnqkgtrtkygrtnqlqaavtslknngiqvgydvvmhkggadgteivnav 120  
 QY 92 YDLIDLGEFNGKGTVRTKYGRSLOQAVTSLKNNGIQVYGDVVMHKGADGTEVMNAV 151  
 Db 121 evnrsnrnqetsgeyaieawtkfdpgrgnhshfkrwryhfdgtdwqsrqlqknykf 180  
 QY 152 EVNRSNRNQEISGEYITAEWTKFDPPGRGNTHSNFKRWYHFDGTDWQSRQLQNKIYKF 211  
 Db 181 rgtgkawdwevdtengnynlymadvdmhpevihelrnwgvvtyntlnldgfridavkh 240  
 QY 212 RGTGKAWDWEVDIENGNDYLMYADIDMDHPEVINELRNWGVWYTNLNDGFRIDAVKH 271  
 Db 241 ikysfrdwlthvrnttgkpmfavaefwkdnlgaieynlktshnsvfdvplhynlyna 300  
 QY 272 IKYSYTRDWLTHVRNTTGKPMFAVEFWKNDLAAIENYLNKTSWNHVSFVDPVLYNLYNA 331  
 Db 301 snsggyvdmrnlngsvvqkphthavtfdvndhdsqgealesfvqgwfplayalvtre 360  
 QY 332 SNSGGYFDMRNLNGSVVQKHPHIAVTFVNDHDSQGEALESFVQSWFKPLAYALILTRE 391  
 Db 361 qgypsvfydygipthgvpamkskidpllqarqtfaigtqhdvfdhddliigwtregns 420  
 QY 392 QGYPSPVFGDYIGIPTHGVPMSKSIDPLLQARQTAYGTQHDYFDHDDIIGWTRGSDS 451  
 Db 421 hpnsglatimsdpggnkmyvgknkagqvrtditgnrtgtvtinadgwnfsvnggvs 480  
 QY 452 HPNSGLATIMSDPGGNKMYVGKHKAGQVRWTDITGNRSGTGTINADGWNFTVNGGAVS 511  
 Db 481 vvvkq 485  
 QY 512 VVWKQ 516

## RESULT 6

ID W12128 standard; protein; 485 AA.  
 AC W12128;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D209N  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;

KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 209 /label= D209N  
 FT W09623873-Al.  
 PN 08-AUG-1996.  
 PD 05-FEB-1996; DK0056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 stability and reduced calcium ion dependency  
 PS Claim 11: 111pp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 variants of the invention were created using site directed, or random,  
 mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 W12142 and W12144 are specifically variants of the alkalophilic bacillus  
 strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 improved thermal stability (such as at temperatures in the range of 40-70  
 degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 dependency. The variants can also have increased alpha-amylolytic  
 activity (especially at pH values in the range of 8.5-10.5), and improved  
 binding of a particular substrate. These variant alpha-amylases also  
 possess improved specificity to a particular substrate, and/or improved  
 specificity with respect to cleavage of substrate. These sequences can  
 be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 beer-making processes. These variants can also be used in the production  
 of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hngtngtmmqyfeywlpndgnhwnrlrddaanlkskgitavwippawkgtsqndvgyga 60  
 QY 32 HHNGTNGTMMQYFEWHLPLNDGNHNRLRDDAANLKSIGITAVWIPPAAWKGTSQNDVGYGA 91  
 Db 61 ydlldlgefnqkgtrtkygrtnqlqaavtslknngiqvgydvvmhkggadgteivnav 120  
 QY 92 YDLIDLGEFNGKGTVRTKYGRSLOQAVTSLKNNGIQVYGDVVMHKGADGTEVMNAV 151  
 Db 121 evnrsnrnqetsgeyaieawtkfdpgrgnhshfkrwryhfdgtdwqsrqlqknykf 180  
 QY 152 EVNRSNRNQEISGEYITAEWTKFDPPGRGNTHSNFKRWYHFDGTDWQSRQLQNKIYKF 211  
 Db 181 rgtgkawdwevdtengnynlymadvdmhpevihelrnwgvvtyntlnldgfridavkh 240  
 QY 212 RGTGKAWDWEVDIENGNDYLMYADIDMDHPEVINELRNWGVWYTNLNDGFRIDAVKH 271  
 Db 241 ikysfrdwlthvrnttgkpmfavaefwkdnlgaieynlktshnsvfdvplhynlyna 300  
 QY 272 IKYSYTRDWLTHVRNTTGKPMFAVEFWKNDLAAIENYLNKTSWNHVSFVDPVLYNLYNA 331  
 Db 301 snsggyvdmrnlngsvvqkphthavtfdvndhdsqgealesfvqgwfplayalvtre 360  
 QY 332 SNSGGYFDMRNLNGSVVQKHPHIAVTFVNDHDSQGEALESFVQSWFKPLAYALILTRE 391  
 Db 361 qgypsvfydygipthgvpamkskidpllqarqtfaigtqhdvfdhddliigwtregns 420  
 QY 392 QGYPSPVFGDYIGIPTHGVPMSKSIDPLLQARQTAYGTQHDYFDHDDIIGWTRGSDS 451  
 Db 421 hpnsglatimsdpggnkmyvgknkagqvrtditgnrtgtvtinadgwnfsvnggvs 480

QY 452 HPNSGLATIMSDPGGKWMYVCKHKGAGQVWRDITGNRSSTVTINADGNGNFTVNGGAVS 511  
 Db 481 vvvkq 485  
 QY 512 VWVQK 516  
 RESULT 7  
 ID W12118 standard; protein; 485 AA.  
 AC W12118;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant K108R  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 OS Sweetener.  
 QS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 108 /label= K108R  
 FT WO9623873-A1.  
 PN 08-AUG-1996.  
 PD 08-AUG-1996.  
 PF 05-FEB-1996; DK0056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; 11lpp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 CC W12142 and W12144 are specifically variants of the alkaphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrate. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;  
 Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 hngtngtmqyfewylpndgnhwnlrddaanlkskgitavwippawkgtsqndvggga 60  
 QY 32 HNGTNGTMQYFEWHLFNDGNHWNLRDDAANLKSKITAVWIPPWAKGTSQNDVGGA 91  
 Db 61 ydlldlgefngkgtvrtkygtrnqlqaavtslknnglqvygvdvmmhrggadgteivnav 120  
 QY 92 YDLIDLGEFNQKGTVRTKYGTRSQLOGAVTSLKNNGIQVYGDVVMHKGADGTEMYNAV 151  
 Db 121 evnrsnrngetgeyaieawtkdfpgrgnhshsfkwrwyhfdgtdwdqgrlqnkykf 180  
 QY 152 EVNRSNRNEISGEYIEAWTKDFPGRGNTHSNFNKRWYHFDGTDWDQGRQKNIYKF 211  
 Db 181 rgtgkdwdevdtenydylymadvdmhpevihelrnwgvvytntlnldgfridavkh 240  
 QY 212 RTGKAWDEVDIENGYDYLYMADYDMDHPEVINELRNKGWVYTYNLNLDGFRIDAVKH 271  
 Db 241 ikysftrdwlthvrtnttgkpmfavaefwkndlgaienylnktswnhsvfdvplhynlyna 300

QY 272 IKSYTRDMLTHVTRNTTGKPMFAVAEFWKNDLAAENTLNKTSWNHSVDFDPLHYNLYNA 331  
 Db 301 snsggyvdmrnlpgsvvqkbpthavtfvdnhdsgpgealesfvcgwkfpalayalvtre 360  
 QY 332 SNSGGYFDMRNLNGSVVQKHPIHAVTFVDNHDSPGGALESFVQSWKFPALAYALLTRE 391  
 Db 361 qgypsvfygdygigipthgvpamkskldpllgartqfaygtghdyfhdhdiqwtregns 420  
 QY 392 QGYPVFGDYIGIPTHGVPMSKSKIDPLLQARQTYAYGTQHDYFDHDIIGWTREGDSS 451  
 Db 421 hpnsglatimsdpgggnkwmvvgknkagqvvrddtngtvtinadgngnfsvnggsvs 480  
 QY 452 HPNSGLATIMSDPGGKNWYVGHKAGQVWRDITGNRSSTVTINADGNGNFTVNGGAVS 511  
 Db 481 vvvkq 485  
 QY 512 VWVQK 516  
 RESULT 8  
 ID W12122 standard; protein; 485 AA.  
 AC W12122;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D163N.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW Sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 163 /label= D163N  
 FT WO9623873-A1.  
 PN 08-AUG-1996.  
 PD 08-AUG-1996.  
 PF 03-FEB-1996; DK0056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; 11lpp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 CC W12142 and W12144 are specifically variants of the alkaphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrate. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;  
 Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 hngtngtmqyfewylpndgnhwnlrddaanlkskgitavwippawkgtsqndvggga 60  
 QY 32 HNGTNGTMQYFEWHLFNDGNHWNLRDDAANLKSKITAVWIPPWAKGTSQNDVGGA 91

Db 61 ydlvdlgefnkgvtrtkytrnglqaaavtslknngiqvgydvmmhkggagdtelnav 120  
 Qy 92 YDLVDLGEFNGKGVTRTKYTRNGLSQQAATSLKNNGIQVGYDVMMHKGAGDGTENVAV 151  
 Db 121 evnsrnnqetsgevalawtkkfipgrgnhshsfkwyhfnngtdwgsqqlnklykf 180  
 Qy 152 EVNSRNNQETSGETTIEATKTFDPGRGNTHSNFKRWYHFDGTDWDSQRLQNKIYKF 211  
 Db 181 rgtkavdwewdtengnydylymadvmdhpevihelrnwgyvwtntlnldgfridavkh 240  
 Qy 212 RGTGKAWDEWDIENGNYDLYMADIDMDHPEVINELRNWGYWTNTLNLDGFRIDAVKH 271  
 Db 241 ikysftrdwthvtrntgkpmfavaefwkndlgaienylnktswnhsvfdvplhynlyna 300  
 Qy 272 IKYSFTRDWLTHVTRNTGKPMFAVAEFWKNDLAAIENYLNKTSWNHSHVDFVPLHYNLYNA 331  
 Db 301 snsggyvdmnrlngsvvqkhphtavtfvdnhdsgpgealesfvqgwfplayalvire 360  
 Qy 332 SNSGGYFDMRNILNGSVVQKHPIHATVFDNHDSPGGEALESFVQSWFKPLAYALITRE 391  
 Db 361 qgypsvfygdygipthgvpamkskidpllqarqtavgtghdyfddhdiigwtregns 420  
 Qy 392 QGYPSVFGYDYGIPTHGVPAMSKIDPLLQARQTYAGTQHDYFDHDDIIGWREGDSS 451  
 Db 421 hpnsglatinsdpggknkmyvgnkagqvrditgnrtgtvtinadgwnfsvngsgvs 480  
 Qy 452 HPNSGLATINS DPGGKNKMYVGNKAGQVWRDITGNRS GTVTINADGWNFTVNGGAVS 511  
 Db 481 vvwkq 485  
 Qy 512 VVWQK 516  
 RESULT 9  
 ID W12123 standard; protein; 485 AA.  
 AC W12123;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D188N.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key  
 FT misc\_difference 188  
 FT /label= D188N  
 PN WO9623873-A1.  
 PD 08-AUG-1996.  
 PF 05-FEB-1996; DK0056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgaard-frantzen H, Borchert T, Svendsen A;  
 DR WPI, 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; il1pp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 CC W12142 and W12144 are specifically variants of the alkaphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrate. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.

CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;

Db 1 hhnngtngtmgyfeywlpndgghwnrlrddaanlkskgtavwippawkgtsqndvggga 60  
 Qy 32 HHNGTNGTMGYFEWHLPNNDGHNHRLRDDAANLKSKGITAYWIPPAPWKGTSQNDVGGA 91  
 Db 61 ydlvdlgefnkgvtrtkytrnglqaaavtslknngiqvgydvmmhkggagdtelnav 120  
 Qy 92 YDLVDLGEFNGKGVTRTKYTRNGLSQQAATSLKNNGIQVGYDVMMHKGAGDGTENVAV 151  
 Db 121 evnsrnnqetsgevalawtkkfipgrgnhshsfkwyhfnngtdwgsqqlnklykf 180  
 Qy 152 EVNSRNNQETSGETTIEATKTFDPGRGNTHSNFKRWYHFDGTDWDSQRLQNKIYKF 211  
 Db 181 rgtkavdwewdtengnydylymadvmdhpevihelrnwgyvwtntlnldgfridavkh 240  
 Qy 212 RGTGKAWDEWDIENGNYDLYMADIDMDHPEVINELRNWGYWTNTLNLDGFRIDAVKH 271  
 Db 241 ikysftrdwthvtrntgkpmfavaefwkndlgaienylnktswnhsvfdvplhynlyna 300  
 Qy 272 IKYSFTRDWLTHVTRNTGKPMFAVAEFWKNDLAAIENYLNKTSWNHSHVDFVPLHYNLYNA 331  
 Db 301 snsggyvdmnrlngsvvqkhphtavtfvdnhdsgpgealesfvqgwfplayalvire 360  
 Qy 332 SNSGGYFDMRNILNGSVVQKHPIHATVFDNHDSPGGEALESFVQSWFKPLAYALITRE 391  
 Db 361 qgypsvfygdygipthgvpamkskidpllqarqtavgtghdyfddhdiigwtregns 420  
 Qy 392 QGYPSVFGYDYGIPTHGVPAMSKIDPLLQARQTYAGTQHDYFDHDDIIGWREGDSS 451  
 Db 421 hpnsglatinsdpggknkmyvgnkagqvrditgnrtgtvtinadgwnfsvngsgvs 480  
 Qy 452 HPNSGLATINS DPGGKNKMYVGNKAGQVWRDITGNRS GTVTINADGWNFTVNGGAVS 511  
 Db 481 vvwkq 485  
 Qy 512 VVWQK 516

RESULT 10  
 ID W12130 standard; protein; 485 AA.  
 AC W12130;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant E194Q.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key  
 FT misc\_difference 194  
 FT /label= E194Q  
 PN WO9623873-A1.  
 PD 08-AUG-1996.  
 PF 05-FEB-1996; DK0056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgaard-frantzen H, Borchert T, Svendsen A;  
 DR WPI, 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; il1pp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The





ID W12127 standard; protein; 485 AA.  
 AC W12127;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D207N.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 207 /label= D207N  
 PN W09623873-A1.  
 PD 08-AUG-1996.  
 PF 05-FEB-1996; DK00056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; 111pp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrates. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 hngtngtmqyfewylpndgnhwnrlrddaanlkskgitavwippawkgtsqndvgyga 60  
 QY 32 HNGTNGTMMQYFEWHLFNDGNHNRLRDDAANLKSIGITAVWIPPWKGTSQNDVGYGA 91  
 Db 61 ydlidylgefndqkgtvrtkygtrnqlqaavtslknngiqvgydvvmnhkkgadgteivnav 120  
 QY 92 YDLIDLGEFNDQKGTVRTKYGTRSQLQGAVTSLKNNGLQVIGDVVNNHKGADGTEMVNAV 151  
 Db 121 evnrsnrnqetgsgeyaleawtkfdpgrgnhshsfkwrwyhfdgtdwdsrqlnkiyxf 180  
 QY 152 EVNRSNRNQEISGEYTTIETAWTKFDPPGRGNTHSNFKRWYHFDGTDWDSRQLNKIYKF 211  
 Db 181 rgtgkdwdevtengnydylymadvnmhpeviheirnvgyvtyntlnldgfridavkh 240  
 QY 212 RGTGKADWDEVDENGNYDYLYMADVNDMDHPEVINEIRNNGVWYNTNLNDGFRIDAVKH 271  
 Db 241 ikysfirdwlthvnrnttqkpmfavaefwkdnlgaieynlnktswnhsvfdvplhynlyna 300  
 QY 272 IKYSYTRDMLTHVRNTTGKPMFAVAEFAEWKNDLAAIENYLNKTSWNHVSFVDVPLHYNLYNA 331  
 Db 301 snsggydmrnlngsvvqkphchavtfvdnhdspqgealesfvqqwfkplayalvltre 360  
 QY 332 SNSGGYFDMRNLNGSVVQKHPHIAVTFVDNHDSPQGEALESFVQSWFKPLAYALILTRE 391  
 Db 361 qgyvpsfygdygyipthgvpamkskldpllqarqtqfaygtqchdyfddhddliigtregns 420

QY 392 QCQPSVFYGDYGIPTHGVPMSKSIDPLQARQTVAYGTQHDYFDHDDIIGWTREGDSS 451  
 Db 421 hpnslatimsdpggnkwmvygknkagqvrdtgnrtgtvtinadgwnfsvngsgvs 480  
 QY 452 HPNSLATIMSDPGGNKWMYVYGKNKAGQVWRDITGNRSCTVTINADGWNFTVNGGAVS 511  
 Db 481 vvwkq 485  
 QY 512 VVWKQ 516  
 RESULT 13  
 ID W12126 standard; protein; 485 AA.  
 AC W12126;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D205N.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 205 /label= D205N  
 PN W09623873-A1.  
 PD 08-AUG-1996.  
 PF 05-FEB-1996; DK00056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; 111pp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC variants of the invention were created using site directed, or random,  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrates, and/or improved  
 CC specificity with respect to cleavage of substrate. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 hngtngtmqyfewylpndgnhwnrlrddaanlkskgitavwippawkgtsqndvgyga 60  
 QY 32 HNGTNGTMMQYFEWHLFNDGNHNRLRDDAANLKSIGITAVWIPPWKGTSQNDVGYGA 91  
 Db 61 ydlidylgefndqkgtvrtkygtrnqlqaavtslknngiqvgydvvmnhkkgadgteivnav 120  
 QY 92 YDLIDLGEFNDQKGTVRTKYGTRSQLQGAVTSLKNNGLQVIGDVVNNHKGADGTEMVNAV 151  
 Db 121 evnrsnrnqetgsgeyaleawtkfdpgrgnhshsfkwrwyhfdgtdwdsrqlnkiyxf 180  
 QY 152 EVNRSNRNQEISGEYTTIETAWTKFDPPGRGNTHSNFKRWYHFDGTDWDSRQLNKIYKF 211

Db 181 rgtgkwdwvntngnydylymyanvdmhpevihelrnwgvvytntlnldgfridavkh 240  
 QY 212 RGTGKAWDEVDIENGNYDYLYMYADIDMDHPEVINELRNWGVVYNTLNLDGFRIDAVKH 271  
 Db 241 ikysfrdwlthvnttggkpmfavaefwknldgalenylntkswnhsvfdvplhynlyna 300  
 QY 272 IKYSYTRDOWLTHVNTTGGKPMFAVAEFWKNDLAAENLYNLTWSNHSVDFVPLHYNLYNA 331  
 Db 301 snsggydmrnlngsvvqkphthavtfdvhdhdsqgealesfvqgwkplayalvltre 360  
 QY 332 SNSSGYFDMRNLNGSVVQKPHTHAVTFVDNHDSDQGEALESFVQSWFKPLAYALILTRE 391  
 Db 361 qgypsvfygdygipthgvpamkskidpllqarqtfaigtqhdhdiigwtregns 420  
 QY 392 QGYPVSFYGDYIGPIPTHGVPAMKSIDPLLQARQTYAYGTQHDYFDHDIIGWTREGDSS 451  
 Db 421 hpnsglatimsdpgggnkwmvqgnkagvvrtdtgnrtgtvtinatdgwnfsvnggsvs 480  
 QY 452 HPNSGLATIMSDPGGKNKWMYVGRKHAGOVVRDITGNRSQVTTINADGWNFTVNGGAVS 511  
 Db 481 vvvkq 485  
 QY 512 VVWKQ 516

RESULT 14  
 ID W12124 standard; protein; 485 AA.  
 AC W12124;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant D192N.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 192 /label= D192N  
 FT W09623873-Al.  
 PN 08-AUG-1996.  
 PD 05-FEB-1996; DK0056.  
 PF 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 11; ; l1lpp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, W12956, W12956, W12956, W12956, W12956,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R18135. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic  
 CC activity (especially at pH values in the range of 8.5-10.5), and improved  
 CC binding of a particular substrate. These variant alpha-amylases also  
 CC possess improved specificity to a particular substrate, and/or improved  
 CC specificity with respect to cleavage of substrate. These sequences can  
 CC be used in detergent and washing compositions, and for textile desizing.  
 CC The alpha-amylase variants can also be used in papermaking and  
 CC beer-making processes. These variants can also be used in the production  
 CC of sweeteners and ethanol from starch.  
 SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
 Best Local Similarity 94.8%; Pred. No. 0.00e+00;  
 Matches 460; Conservative 19; Mismatches 6; Indels 0; Caps 0;

Db 1 hngtngtmqyfwylypndgnhwnrlrddaanlkskgtitavwippawksqsdvgyga 60  
 QY 32 HHNGTNGTMQYFEWHLPLNDGNEHNRLLRDAANLKSKITAVWIPPAAWKGSQNDVGYGA 91  
 Db 61 ydyldlqefnqktvrtkytrnqlgaavtslknngigvygdyvnmhkhggadgteivnav 120  
 QY 92 YDYLDLGEFNQKTVRTKYTRNQLGAAVTSLKNNGIQVYGDVVMHKGADGTEVMNAV 151  
 Db 121 evnrsnrgtsgyaleawtkfdpgrgnhshsfkwrwyhfdgtwdqsrqlgnkiyxf 180  
 QY 152 EVNRSNRNQBISGEYIEAWTKFDPGRGNTHSNFKRWYHFDCTDWDQSRQLGNKIYKF 211  
 Db 181 rgtgkwdwvntngnydylymyadidmdhpevihelrnwgvvytntlnldgfridavkh 240  
 QY 212 RGTGKAWDEVDIENGNYDYLYMYADIDMDHPEVINELRNWGVVYNTLNLDGFRIDAVKH 271  
 Db 241 ikysfrdwlthvnttggkpmfavaefwknldgalenylntkswnhsvfdvplhynlyna 300  
 QY 272 IKYSYTRDOWLTHVNTTGGKPMFAVAEFWKNDLAAENLYNLTWSNHSVDFVPLHYNLYNA 331  
 Db 301 snsggydmrnlngsvvqkphthavtfdvhdhdsqgealesfvqgwkplayalvltre 360  
 QY 332 SNSSGYFDMRNLNGSVVQKPHTHAVTFVDNHDSDQGEALESFVQSWFKPLAYALILTRE 391  
 Db 361 qgypsvfygdygipthgvpamkskidpllqarqtfaigtqhdhdiigwtregns 420  
 QY 392 QGYPVSFYGDYIGPIPTHGVPAMKSIDPLLQARQTYAYGTQHDYFDHDIIGWTREGDSS 451  
 Db 421 hpnsglatimsdpgggnkwmvqgnkagvvrtdtgnrtgtvtinatdgwnfsvnggsvs 480  
 QY 452 HPNSGLATIMSDPGGKNKWMYVGRKHAGOVVRDITGNRSQVTTINADGWNFTVNGGAVS 511  
 Db 481 vvvkq 485  
 QY 512 VVWKQ 516

RESULT 15  
 ID W12108 standard; protein; 485 AA.  
 AC W12108;  
 DT 08-APR-1997 (first entry)  
 DE Alpha-amylase variant K269R.  
 KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
 KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
 KW textile desizing; papermaking; beer-making; ethanol production;  
 KW sweetener.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 269 /label= K269R  
 FT W09623873-Al.  
 PN 08-AUG-1996.  
 PD 05-FEB-1996; DK0056.  
 PR 03-FEB-1995; DK-000126.  
 PR 29-MAR-1995; DK-000336.  
 PR 29-SEP-1995; DK-001097.  
 PR 06-OCT-1995; DK-001121.  
 PA (NOVO ) NOVO-NORDISK AS.  
 PI Bisgard-frantzen H, Borchert T, Svendsen A;  
 DR WPI; 96-371423/37.  
 PT Alpha-amylase variants - with improved thermal and oxidation  
 PT stability and reduced calcium ion dependency  
 PS Claim 10; ; l1lpp; English.  
 CC W12098-W12144 represent alpha-amylase variants of the invention. The  
 CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
 CC represented by W12955, W12956, W12956, W12956, W12956, W12956,  
 CC W12142 and W12144 are specifically variants of the alkaliphilic Bacillus  
 CC strain NCIB 12512 alpha-amylase shown in R18135. These variants can have  
 CC improved thermal stability (such as at temperatures in the range of 40-70  
 CC degrees Celsius), and/or oxidation stability, and/or reduced calcium ion  
 CC dependency. The variants can also have increased alpha-amylolytic

CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 485 AA;

Query Match 91.3%; Score 3536; DB 20; Length 485;  
Best Local Similarity 94.8%; Pred. NO. 0.00e+00;  
Matches 460; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
Db 1 hhnngntmmgyfeywlpndgnhwnrlrdaanlkskgitavwipawkgtsqndvgyga 60  
Qy 32 HNGTNGTMMQYFEWHPNDGNHWNRLRDDAANLKSKITAVWIPAWKTSQNDVGYGA 91  
Db 61 ydlldlgefnqgtvrtkygtrnqlqaavtslknngigvygvdvmmhkggadgteivnav 120  
Qy 92 YDLYDLGEFNGQGTVRTKYGTRSQLQGA VTS LKNNGIQVYGVVMMHKGADGTEMVNAV 151  
Db 121 evnrsnrnqetsgeyaieawtkfdipgrgnhssfkrrwyhfdgtwdqsrqlqnkiykf 180  
Qy 152 EVNRSNRNQEISGEYIEAWTKFDIPGRGNTHSFKRWYHFDGTDWDSRQLQNKIYKF 211  
Db 181 rgtkawkdewdtengnydylmadvdmhpevihelrnwgvwytnlndgfridavkh 240  
Qy 212 RGTGAWDWEVDIENGNDYLMYADIDMDHPEVINELRNWGVWYTNLNDGFRIDAVKH 271  
Db 241 lkysfrdlwltvhrnttgkpmfavaefwrndlgalenylnktswnhsvfdvplhynlyna 300  
Qy 272 IKYSYTRDOWLTHVRNTTGKPMFAVAEFWKNDLAAIENYLNKTSWNHSVFDVPLHYNLNA 331  
Db 301 snsggydmrnlningsvqkphthavtfvnhdsqgealesfvqgwkfplayalvtre 360  
Qy 332 SNSGGYFDMRNILNGSVQKHPIHAVTFVDNHDSPQGEALESFVQSWFKPLAYALILTRE 391  
Db 361 ggypsvfgydygipthgypamksklpdlqarqtfaqtdydhhdliigwtregns 420  
Qy 392 QGYPSVFYGDYGIPTHGYPMSKSIDPLQLARQTYAGTQHDYFDHHDIIIGWTREGDSS 451  
Db 421 hpnsqglatimsdpgggnkwmvygknkagqvrdrditgnrtgtvtinadgwnfsvngsvs 480  
Qy 452 HPNSGLATIMSDPGGNKWMYVGKKAGQVWRDITGNRSQTVTINADGWNFTVNGGAVS 511  
Db 481 vvvkq 485  
Qy 512 VVWKQ 516

Search completed: Tue Sep 15 14:16:18 1998  
Job time : 150 secs.

\*\*\*\*\*

WIREHOLE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Sep 15 14:22:04 1998; MasPar time 7.52 Seconds  
Tabular output not generated. 484.408 Million cell updates/sec

Title: >US-08-952-741-2  
Description: (1-516) from US08952741.pep  
Perfect Score: 3873  
Sequence: 1 MKLHNRIISVLLTLLAVV.....ADGNGFTVNGAVSVWYKQ 516

Scoring table: PAM 150  
Gap 11

Searched: 77021 seqs, 7058996 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 33.562; Variance 161.785; scale 0.207

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description                       | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|
| 1          | 2682  | 69.2        | 512    | 1  | US-07-623- Sequence 5, Applicatio | 6.31e-224 |
| 2          | 2680  | 69.2        | 512    | 1  | US-08-720- Sequence 2, Applicatio | 9.51e-224 |
| 3          | 2671  | 69.0        | 511    | 1  | US-08-468- Sequence 35, Applicati | 6.03e-223 |
| 4          | 2671  | 69.0        | 511    | 1  | US-08-645- Sequence 3, Applicatio | 6.03e-223 |
| 5          | 2671  | 69.0        | 511    | 2  | PCT-US94-0 Sequence 33, Applicati | 6.03e-223 |
| 6          | 2671  | 69.0        | 511    | 2  | PCT-US95-1 Sequence 33, Applicati | 6.03e-223 |
| 7          | 2662  | 68.7        | 512    | 1  | US-07-623- Sequence 3, Applicatio | 4.69e-222 |
| 8          | 2661  | 68.7        | 512    | 1  | US-08-720- Sequence 6, Applicatio | 1.80e-219 |
| 9          | 2632  | 68.0        | 483    | 2  | PCT-US95-1 Sequence 32, Applicati | 1.80e-219 |
| 10         | 2632  | 68.0        | 483    | 2  | PCT-US94-0 Sequence 32, Applicati | 1.80e-219 |
| 11         | 2632  | 68.0        | 483    | 1  | US-08-468- Sequence 34, Applicati | 1.80e-219 |
| 12         | 2632  | 68.0        | 487    | 2  | PCT-US95-1 Sequence 37, Applicati | 1.80e-219 |
| 13         | 2632  | 68.0        | 487    | 2  | PCT-US94-0 Sequence 37, Applicati | 1.80e-219 |
| 14         | 2629  | 67.9        | 483    | 1  | US-08-645- Sequence 2, Applicatio | 3.34e-219 |
| 15         | 2620  | 67.6        | 483    | 2  | PCT-US95-1 Sequence 36, Applicati | 2.11e-218 |
| 16         | 2620  | 67.6        | 483    | 2  | PCT-US94-0 Sequence 36, Applicati | 2.11e-218 |
| 17         | 2599  | 66.1        | 514    | 1  | US-08-720- Sequence 4, Applicatio | 5.75e-213 |
| 18         | 2599  | 66.1        | 520    | 1  | US-08-645- Sequence 4, Applicatio | 5.75e-213 |
| 19         | 2599  | 66.1        | 520    | 1  | US-08-468- Sequence 36, Applicati | 5.75e-213 |
| 20         | 2599  | 66.1        | 520    | 2  | PCT-US95-1 Sequence 34, Applicati | 5.75e-213 |
| 21         | 2599  | 66.1        | 520    | 2  | PCT-US94-0 Sequence 34, Applicati | 5.75e-213 |
| 22         | 2524  | 65.2        | 548    | 2  | PCT-US94-0 Sequence 35, Applicati | 7.54e-210 |
| 23         | 2524  | 65.2        | 548    | 1  | US-08-645- Sequence 5, Applicatio | 7.54e-210 |

|    |      |      |     |   |            |                        |           |
|----|------|------|-----|---|------------|------------------------|-----------|
| 24 | 2524 | 65.2 | 548 | 2 | PCT-US95-1 | Sequence 35, Applicati | 7.54e-210 |
| 25 | 2524 | 65.2 | 548 | 1 | US-08-468- | Sequence 37, Applicati | 7.54e-210 |
| 26 | 1899 | 49.0 | 484 | 3 | 5171673-8  | Patent No. 5171673.    | 2.83e-154 |
| 27 | 1899 | 49.0 | 487 | 3 | 5171673-6  | Patent No. 5171673.    | 2.83e-154 |
| 28 | 284  | 7.3  | 60  | 3 | 5171673-4  | Patent No. 5171673.    | 1.18e-13  |
| 29 | 174  | 4.5  | 38  | 3 | 5171673-2  | Patent No. 5171673.    | 5.53e-05  |
| 30 | 138  | 3.6  | 428 | 1 | US-08-343- | Sequence 4, Applicatio | 2.55e-02  |
| 31 | 138  | 3.6  | 428 | 1 | US-07-973- | Sequence 4, Applicatio | 2.55e-02  |
| 32 | 138  | 3.6  | 434 | 1 | US-08-184- | Sequence 6, Applicatio | 2.55e-02  |
| 33 | 138  | 3.6  | 434 | 1 | US-07-923- | Sequence 6, Applicatio | 2.55e-02  |
| 34 | 135  | 3.5  | 687 | 1 | US-08-470- | Sequence 10, Applicati | 4.20e-02  |
| 35 | 135  | 3.5  | 687 | 1 | US-08-467- | Sequence 10, Applicati | 4.20e-02  |
| 36 | 132  | 3.4  | 687 | 1 | US-08-204- | Sequence 10, Applicati | 6.87e-02  |
| 37 | 127  | 3.3  | 478 | 1 | US-08-720- | Sequence 7, Applicatio | 1.56e-01  |
| 38 | 124  | 3.2  | 437 | 1 | US-07-973- | Sequence 6, Applicatio | 2.53e-01  |
| 39 | 124  | 3.2  | 437 | 1 | US-08-343- | Sequence 6, Applicatio | 2.53e-01  |
| 40 | 120  | 3.1  | 438 | 1 | US-07-973- | Sequence 2, Applicatio | 4.81e-01  |
| 41 | 120  | 3.1  | 438 | 1 | US-08-343- | Sequence 2, Applicatio | 4.81e-01  |
| 42 | 119  | 3.1  | 468 | 1 | US-08-470- | Sequence 9, Applicatio | 5.65e-01  |
| 43 | 119  | 3.1  | 468 | 1 | US-08-204- | Sequence 8, Applicatio | 5.65e-01  |
| 44 | 119  | 3.1  | 468 | 1 | US-08-467- | Sequence 9, Applicatio | 5.65e-01  |
| 45 | 114  | 2.9  | 468 | 1 | US-08-467- | Sequence 6, Applicatio | 1.25e+00  |

ALIGNMENTS

|          |                 |           |      |         |
|----------|-----------------|-----------|------|---------|
| RESULT 1 | US-07-623-953-5 | STANDARD; | PRT; | 512 AA. |
| ID       | XX              | XX        | XX   | XXXXXX  |
| AC       | XX              | XX        | XX   | XXXXXX  |
| DT       | XX              | XX        | XX   | XXXXXX  |

Sequence 5, Application US/07623953

Sequence 5, Application US/07623953

Patent No. 5364782

GENERAL INFORMATION:

APPLICANT: Quax, Wilhelmus J

APPLICANT: Laroche, Yves

APPLICANT: Vollebregt, Adrianus W.H.

APPLICANT: Stanssens, Patrick

APPLICANT: Lauwereys, Marc

TITLE OF INVENTION: MUTANT MICROBIAL ALPHA-AMYLASES WITH

TITLE OF INVENTION: INCREASED THERMAL, ACID AND/OR ALKALINE STABILITY

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM

STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR

CITY: PALO ALTO

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/07/623,953

APPLICATION NUMBER: 19901129

FILING DATE: 19901129

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rae-Venter, Barbara

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: GBRO-025/0005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-494-7622

TELEFAX: 415-857-0663

TELEX: 380816 COOLEY PA

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 512 amino acids

Nov 15, 94







```

QY 361 DNHDSQPGEALSFVQSFKPLAYALILITREGQYPSVFGDYGI--PTH-GVPSMKSKI 417
Db 413 EPIKARKQYAGAOHDYFDHHDIVGWTREGDSSVANSGLAALITDGPQAKRMVVGRO 472
QY 418 DPLQARQYAYGTQHDYFDHHDIVGWTREGDSSHPNSGLATINSDGPGGNKMYVGRHK 477
Db 473 AGETHDITGNRSEPVVINSSEGEFHVHNGSVSIYVOR 511
QY 478 AGQVWRDITGNRSGTGTINADGNGFTVNGGAVSVWVKQ 516

RESULT 5
ID PCT-US94-01553A-33 STANDARD; PRT; 511 AA.
XX
AC xxxxxx
XX
DT
XX
XX
DE DE
CC Sequence 33, Application PC/TUS9401553A
CC Sequence 33, Application PC/TUS9401553A
CC GENERAL INFORMATION:
CC APPLICANT: GENENCOR INTERNATIONAL, INC.
CC TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase
CC NUMBER OF SEQUENCES: 68
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genencor International, Inc.
CC STREET: 180 Kimball Way
CC CITY: South San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94080
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/01553A
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hoin, Margaret A.
CC REGISTRATION NUMBER: 33,401
CC REFERENCE/DOCKET NUMBER: GC220-2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 742-7536
CC TELEFAX: (415) 742-7217
CC INFORMATION FOR SEQ ID NO: 33:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 511 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 511 AA; 58364 MW; 1365410 CN;

Query Match 69.0%; Score 2671; DB 2; Length 511;
Best Local Similarity 66.9%; Pred. No. 6.03e-223;
Matches 347; Conservative 86; Mismatches 75; Indels 11; Gaps 8;

Db 1 MKQQRRLVRLTLFLALIFLPH-SAAAA--N-INGTLMQYFEWMPNDG-HWKRLQN 55
QY 1 MKNHRIISVLLTLLAVLFPYMTPEAQAHHNGTNGTMMQYFEWHLFNDGNHNRLED 60
Db 56 DSAYLAEHGITAVTPPAKGTQADYGVAYDLDLGEFHQKGVRYKYGKELQSAI 115
QY 61 DAANKSGITAVTPPAKGTQADYGVAYDLDLGEFNQKGVRYKYGTRSLQGVAV 120
Db 116 KSLSRSDINVTGDDVYVNHKGGADATEDVTAVEVDPADNRNVSGEHLKAWTHFHPGRG 175
QY 121 TSLKNGNQVYGDVVMNHKGGADGTEMVNAVENVNRRNQELSGEYTTAEWTKDFPGRG 180

```

```

Db 176 STYSDFKHWHFDTGTDWDESKL-NRIYKFOG--KAWDWEVSNENGNNDYILMYADIDYD 232
QY 181 NTHSNFKRWTHFDTGTDWDESKL-NRIYKFOG--KAWDWEVSNENGNNDYILMYADIDYD 240
Db 233 HPDVAAEIKRMTWYANELQLDGFRLDAVKHIFKSFRLDWNHVRKTKGKEMFTVAEYHQ 292
QY 241 HPEVINELRNMGVWYNTNLNDGFRIDAVKHIFKSYTRDLWLTHTVRNTTCKPMPFAVEFWK 300
Db 293 NDLGALENYLNKTNFNHNSVDFVPLHYQFHAASSTOGGDMRKLNGTVVSKHPLKSVTFV 352
QY 301 NDAAIENYLNKTSWNHNSVDFVPLHYNLYNASNGGYEDMRNLNGSVVQKHPHIAVTFV 360
Db 353 DNHDTPQGSLESTVOTWFKPLAYAFILITRESGYPOVFGDMYGTGDSQRIKPKHKI 412
QY 361 DNHDSQPGEALSFVQSFKPLAYALILITREGQYPSVFGDYGI--PTH-GVPSMKSKI 417
Db 413 EPIKARKQYAGAOHDYFDHHDIVGWTREGDSSVANSGLAALITDGPQAKRMVVGRO 472
QY 418 DPLQARQYAYGTQHDYFDHHDIVGWTREGDSSHPNSGLATINSDGPGGNKMYVGRHK 477
Db 473 AGETHDITGNRSEPVVINSSEGEFHVHNGSVSIYVOR 511
QY 478 AGQVWRDITGNRSGTGTINADGNGFTVNGGAVSVWVKQ 516

RESULT 6
ID PCT-US95-10426-33 STANDARD; PRT; 511 AA.
XX
AC xxxxxx
XX
DT
XX
XX
DE DE
CC Sequence 33, Application PC/TUS9510426
CC Sequence 33, Application PC/TUS9510426
CC GENERAL INFORMATION:
CC APPLICANT: GENENCOR INTERNATIONAL, INC.
CC TITLE OF INVENTION: An Improved Cleaning Composition
CC NUMBER OF SEQUENCES: 68
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genencor International
CC STREET: 180 Kimball Way
CC CITY: South San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94080
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/10426
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: STONE, Christopher L.
CC REGISTRATION NUMBER: 33,401
CC REFERENCE/DOCKET NUMBER: GC220-3
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 742-7536
CC TELEFAX: (415) 742-7217
CC INFORMATION FOR SEQ ID NO: 33:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 511 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 511 AA; 58364 MW; 1365410 CN;

Query Match 69.0%; Score 2671; DB 2; Length 511;
Best Local Similarity 66.9%; Pred. No. 6.03e-223;

```



CC APPLICANT: Thellersen, Marianne  
CC APPLICANT: Van der Zee, Pia  
CC TITLE OF INVENTION: AMYLASE VARIANTS  
CC NUMBER OF SEQUENCES: 38  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: No. 57534600 No. 57534600disk of No. 5753460th America, Inc.  
CC STREET: 405 Lexington Avenue, 64th Floor  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10174-6401  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/720,899  
CC FILING DATE: 10-OCT-1996  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/343,804  
CC FILING DATE: 22-NOV-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Lowney Dr. Karen A.  
CC REGISTRATION NUMBER: 31,274  
CC REFERENCE/DOCKET NUMBER: 4054, 214-US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 212-867-0123  
CC TELEFAX: 212-878-9655  
CC INFORMATION FOR SEQ ID NO: 6:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 549 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 549 AA; 62566 MW; 1687311 CN;

Query Match 68.7%; Score 2661; DB 1; Length 549;  
Best Local Similarity 67.5%; Pred. No. 4.69e-222;  
Matches 341; Conservative 78; Mismatches 81; Indels 5; Gaps 5;

Db 15 LLAPLLTVS-LFCPTGPAPAA-APEFNGTMMQYFEWLPDDGTMTKVNANLSSLGI 72  
Qy 11 LLLTLLAVAVLPYMTPEPAQHNGTNGTMMQYFEWLPDNGNHNRLRDAANLKSNGI 70  
Db 73 TALWLPAYKGTSDSYGYDYLDLGEFNGKGTVRTKGTQYLYQ-AIQAAHAAGMQ 131  
Qy 71 TAVWIPPAWKGTSDQYGYGAYLDLGEFNGKGTVRTKGTQYLYQ-LOGAVTSLKNGIQ 129  
Db 132 VYADVVDHKGAGDGTWVDVAVEVNPDRNQEISGTQIQAWTKFDPGRGNTYSSFKWR 191  
Qy 130 YGDDVNNHKGAGDGTWVNAVEVNNRNOEISGEYTIKAWTKFDPGRGNTSNEFKWR 189  
Db 192 WYHFDGVDWDESRKL-SRIYKFRIGKAWDWEVDTEGNDYDLYMADMDHPVVTLEK 250  
Qy 190 WYHFDGTDQWQSLQNLQKIIYKFGTGAWDWEVDIENGNDYDLYMADMDHPVINELR 249  
Db 251 NWGKVVNTNIDGFRDLDAVKKHKSFPDWSYRSQTKPLFTVGEYWSYDINKLHNY 310  
Qy 250 NWGKVVNTNLDGFRDLDAVKKHKSYSYTRDWTGKPMFAVEAFKNDLAAATENY 309  
Db 311 ITKDTGMSLFDAPLNKKEYTASKSGAFDMRVLTMTLTKMDQPTLATVFDVNDHDEPGQ 370  
Qy 310 LKNTSNHNSVDFVPLHNLNNAISGGYFDMRNILNNGSVQKHPHIAVTVDVNDHSQPE 369  
Db 371 ALOSWDPMFKPLAYAFLTRQEGYCVFYGYGIPQYINPSLKSIDPLLIARDYAY 430  
Qy 370 ALESFVQSKFKPLAYALITREQGYSVFYGYGIPTHGVPSMKSIDPLLIARQYAY 429  
Db 431 GTQHDYLDHSDIIGWTREGTEKPGSGLAALITDGGGSKWMTYVQKHAKVFDYLTGNNR 490  
Qy 430 GTQHDYFDHDDIIGWTREGDSSHPNSGLATIMSDGPGGNKWMYVYGHKAGQVYRDTGNNR 489

Db 491 SDVTINSDGGEFKYNGSGSVVWV 515  
Qy 490 SGTVTINADGNGFTYNGGAVSVWV 514

RESULT 9  
ID PCT-US95-10426-32 STANDARD: PRT: 483 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
XX  
DE Sequence 32, Application PC/TUS9510426  
XX  
DE Sequence 32, Application PC/TUS9510426  
CC GENERAL INFORMATION:  
CC APPLICANT: GENENCOR INTERNATIONAL, INC.  
CC TITLE OF INVENTION: An Improved Cleaning Composition  
CC NUMBER OF SEQUENCES: 68  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Genencor International  
CC STREET: 180 Kimball Way  
CC CITY: South San Francisco  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 94080  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/10426  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: STONE, Christopher L.  
CC REGISTRATION NUMBER: 33,401  
CC REFERENCE/DOCKET NUMBER: GC220-3  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 742-7536  
CC TELEFAX: (415) 742-7217  
CC INFORMATION FOR SEQ ID NO: 32:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 483 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 483 AA; 55211 MW; 1221290 CN;

Query Match 68.0%; Score 2632; DB 2; Length 483;  
Best Local Similarity 68.9%; Pred. No. 1.80e-219;  
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;

Db 4 NGTLMQYFWMYNDGQHWKRLONDSAYLAHGHTAVWIPPAYKGTSDQYGYGAYDLYD 63  
Qy 37 NGTMMQYFEWLPDNGNHNRLRDAANLKSNGITAVWIPPAYKGTSDQYGYGAYDLYD 96  
Db 64 LGFHFQKGTVRTKGTGELQSAIKSLHSRDIYNGVDVVIHKKGADATEDVTAVEVDP 123  
Qy 97 LGFHNKGTVRTKGTGELQSAIKSLHSRDIYNGVDVVIHKKGADATEDVTAVEVDP 156  
Db 124 DRNRVISGEHLIKAWTHFHPGSGTYSDFKWHYHFDGTDWDESRKL-NRIYKFG--K 180  
Qy 157 NRQETSGEYTIKAWTKFDPGRGNTSNEFKRWYHFDGTDWDESRKL-NRIYKFG--K 216  
Db 181 AWDWEVSNNGNDYDLYMADMDHPVAAEIKRWGTYWYANELQDGFRLDAVKHIFSF 240  
Qy 217 AWDWEVDIENGNDYDLYMADMDHPVINELRNMGWYNTNLDGFRIDAVKHIXSY 276  
Db 241 LRDWVNHVREKTKEMETVAYEQNDLGALENYLNKTNFNHNSVEDVPLHYQFHAASQTGG 300

|                       |  |  |                   |
|-----------------------|--|--|-------------------|
| QY                    | 277  | TRDLWLRHVRNTTGRKPFVAEFAEWNKDLAAIENLYLNKTSWNHVSFVDPPLHYNLYNASNSGG | 336               |
| Db                    | 301  | GYDMRKLLNGTVVYSKPKLSVTEVDNHDTPQGOSLESTVQTWFKPLAYAFILTRREGGYQ     | 360               |
| QY                    | 337  | YFDMRNILNGSVVQKPIHAVTEVDNHDSDQPGALESFVQSWFKPLAYALLITREGGYPS      | 396               |
| Db                    | 361  | VFYGDMYGTGDSQREIPALKHKIEPIELKARKOYAYGAQHDYFDHHDIVGWTREGDSSVA     | 420               |
| QY                    | 397  | VFYGDYGI--PTH-GVPSMKSKRIDPLLOARQTYAYCTQHDYFDHHDIIIGWTREGDSSHP    | 453               |
| Db                    | 421  | NSGLAALITDPCGAKRMVYGRONACETHWHDITGNSEPEVWINSEGCWGFHVNGGSVSIY     | 480               |
| QY                    | 454  | NSGLATIMSDPGGNKWMVYGVKKHKGQVWRHDITGNRSRGTVTINADGWGNTVNGGAVSW     | 513               |
| Db                    | 481  | VQR 483  |                   |
| QY                    | 514  | VQK 516  |                   |
| RESULT                | 10   |  |                   |
| ID                    | PCT-US94-01553A-32                                   | STANDARD; PRT; 483 AA.   |                   |
| XX                    | xxxxxx   |  |                   |
| XX                    |  |  |                   |
| DT                    |  |  |                   |
| XX                    |  |  |                   |
| DE                    |  |  |                   |
| XX                    |  |  |                   |
| CC                    | Sequence 32, Application PC/TUS9401553A              |  |                   |
| CC                    | Sequence 32, Application PC/TUS9401553A              |  |                   |
| CC                    | GENERAL INFORMATION:                                 |  |                   |
| CC                    | APPLICANT: GENENCOR INTERNATIONAL, INC.              |  |                   |
| CC                    | TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase |  |                   |
| CC                    | NUMBER OF SEQUENCES: 68                              |  |                   |
| CC                    | CORRESPONDENCE ADDRESS:                              |  |                   |
| CC                    | ADDRESSEE: Genencor International, Inc.              |  |                   |
| CC                    | STREET: 180 Kimball Way                              |  |                   |
| CC                    | CITY: South San Francisco                            |  |                   |
| CC                    | STATE: CA  |  |                   |
| CC                    | COUNTRY: USA   |  |                   |
| CC                    | ZIP: 94080   |  |                   |
| CC                    | COMPUTER READABLE FORM:                              |  |                   |
| CC                    | MEDIUM TYPE: Floppy disk                             |  |                   |
| CC                    | COMPUTER: IBM PC compatible                          |  |                   |
| CC                    | OPERATING SYSTEM: PC-DOS/MS-DOS                      |  |                   |
| CC                    | SOFTWARE: Patent In Release #1.0, Version #1.25      |  |                   |
| CC                    | CURRENT APPLICATION DATA:                            |  |                   |
| CC                    | APPLICATION NUMBER: PCT/US94/01553A                  |  |                   |
| CC                    | FILING DATE:   |  |                   |
| CC                    | CLASSIFICATION:                                      |  |                   |
| CC                    | ATTORNEY/AGENT INFORMATION:                          |  |                   |
| CC                    | NAME: Horn, Margaret A.                              |  |                   |
| CC                    | REGISTRATION NUMBER: 33,401                          |  |                   |
| CC                    | REFERENCE/DOCKET NUMBER: GC220-2                     |  |                   |
| CC                    | TELECOMMUNICATION INFORMATION:                       |  |                   |
| CC                    | TELEPHONE: (415) 742-7536                            |  |                   |
| CC                    | TELEFAX: (415) 742-7217                              |  |                   |
| CC                    | INFORMATION FOR SEQ ID NO: 32:                       |  |                   |
| CC                    | SEQUENCE CHARACTERISTICS:                            |  |                   |
| CC                    | LENGTH: 483 amino acids                              |  |                   |
| CC                    | TYPE: amino acid                                     |  |                   |
| CC                    | STRANDEDNESS: single                                 |  |                   |
| CC                    | TOPOLOGY: linear                                     |  |                   |
| CC                    | MOLECULE TYPE: protein                               |  |                   |
| CC                    | SEQUENCE 483 AA; 55211 MW; 1221290 CN;               |  |                   |
| Query Match           | 68.0%;   | Score 2632;  | DB 2; Length 483; |
| Best Local Similarity | 68.9%;   | Pred. No. 1.80e-219;   |                   |
| Matches 333;          | Conservative 77;                                     | Mismatches 67;   | Indels 6; Gaps    |
| Db                    | 4  | NGTLMQYFEWMPNDGQHWRLRQNDLSAYLAEHGITAWIPPAYKGTQSQADVGAYDLYD       | 63                |
| QY                    | 37   | NGTMMQYFEWHLRNDGNHNRRLRDRDANLKSKGITAWIPPAWKGTQSDNDVGAYDLYD       | 96                |

|        |  |   |              |
|--------|--|---|--------------|
| Db     | 64   | LGSEHOKGTVRTYKTKGELQSAIKSLHSDINDINVGDVVINKHGGADATEDVTAVEVDA     | 121          |
| Qy     | 97   | LGSEFNOKGTVRTYKTKGTRSLQQAIVTSLKNNGIQVYGDVVNNHKKGGADGTEMVNAVEVNR | 156          |
| Db     | 124  | DRNRVSGEHLIKAWTHFHPGPGSGTYSDFKWHMYHFDGTDWDDESRKL-NRIYKFOG--K    | 180          |
| Qy     | 157  | NRQOISGETTICAWTKFDPFGGNTHSNFKWRWYHFDGTDWDQSQQLQNKIYKFRGTGK      | 216          |
| Db     | 181  | AWDEVSENGNENTDYLMIADIDYDHPDVAEIKRWGTWTYANELLQDLGFRIDAVKHKFSF    | 240          |
| Qy     | 217  | AWDEWVDIENGNDYLMYADIDMDHPEVINELRNWGWYVNTLNLDOGRIDAVKHKIKSY      | 276          |
| Db     | 241  | LROWNVHREKTKEMPTVAEYQNDLGALENLYLNTKNTFNHNSVDFVPLHYQFHAASTQGG    | 300          |
| Qy     | 277  | TROWLTHVRNNTGKPFAPFAVEFKWDLAAIENTLYLNTKTSWNHNSVDFVPLHYLNNASNG   | 336          |
| Db     | 301  | GYDMRKLNGTVAWTKPLKSVTFVDNHDTPQGOSLESTVQTFWKPLAYAFILTRREGGYPC    | 360          |
| Qy     | 337  | YFDMRNILNSVQKHPIHAVTFVDNHDSDPGGALSFVQSWFKPLAYALILTRREGGYPS      | 396          |
| Db     | 361  | VFYGDYGTGKDSQREIPALKHKEIPILKARKQOYAYGAOHDYFDHHDIIIGWTRREGDSSVA  | 420          |
| Qy     | 397  | VFYGDYGI--PTH-GVPSMKSIDPLLOARQTYAYGTQHDYFDHHDIIIGWTRREGDSSHP    | 453          |
| Db     | 421  | NSGLAALITDGGPGAKRMVYGRONRAGETHWIDITGNRSEPVVINSEGWGEFHVNGGSVSIY  | 480          |
| Qy     | 454  | NSGLATIMSDPGGNKWMYVYKHKAGQVYWRDITGNRSCTVTINADWGNFTVNGGAVSVW     | 513          |
| Db     | 481  | VOR 483   |              |
| Qy     | 514  | VQK 516   |              |
| RESULT | 11   |   |              |
| ID     | US-08-468-700-34                                     | STANDARD;   | PRT; 483 AA. |
| XX     | XXXXX  |   |              |
| AC     |  |   |              |
| XX     |  |   |              |
| DT     |  |   |              |
| XX     |  |   |              |
| DE     |  |   |              |
| XX     |  |   |              |
| CC     | Sequence 34, Application US/08468700                 |   |              |
| CC     | Sequence 34, Application US/08468700                 |   |              |
| CC     | Patent No. 5736499                                   |   |              |
| CC     | GENERAL INFORMATION:                                 |   |              |
| CC     | APPLICANT: COLIN MITCHINSON                          |   |              |
| CC     | APPLICANT: CAROL A. REQUADT                          |   |              |
| CC     | APPLICANT: TRACI H. ROPE                             |   |              |
| CC     | APPLICANT: LEIF P. SOLHEIM                           |   |              |
| CC     | TITLE OF INVENTION: MUTANT ALPHA-AMYLASE             |   |              |
| CC     | NUMBER OF SEQUENCES: 40                              |   |              |
| CC     | CORRESPONDENCE ADDRESS:                              |   |              |
| CC     | ADDRESSEE: Genecor International                     |   |              |
| CC     | STREET: 180 Kimball Way                              |   |              |
| CC     | CITY: South San Francisco                            |   |              |
| CC     | STATE: CA  |   |              |
| CC     | COUNTRY: USA   |   |              |
| CC     | ZIP: 94080   |   |              |
| CC     | COMPUTER READABLE FORM:                              |   |              |
| CC     | MEDIUM TYPE: Floppy disk                             |   |              |
| CC     | COMPUTER: IBM PC compatible                          |   |              |
| CC     | OPERATING SYSTEM: PC-DOS/MS-DOS                      |   |              |
| CC     | SOFTWARE: patentin Release #1.0, Version #1.25 (EPO) |   |              |
| CC     | CURRENT APPLICATION DATA:                            |   |              |
| CC     | APPLICATION NUMBER: US/08/468,700                    |   |              |
| CC     | FILING DATE: 06-JUN-1995                             |   |              |
| CC     | CLASSIFICATION: C52                                  |   |              |
| CC     | ATTORNEY/AGENT INFORMATION:                          |   |              |
| CC     | NAME: Stone, Christopher                             |   |              |
| CC     | REGISTRATION NUMBER: 35,696                          |   |              |
| CC     | REFERENCE/DOCKET NUMBER: GC275                       |   |              |
| CC     | TELECOMMUNICATION INFORMATION:                       |   |              |

```

Sequence 34, Application US/08468700
Patent No. 5736499
GENERAL INFORMATION:
  APPLICANT: COLIN MITCHINSON
  APPLICANT: CAROL A. REQUADT
  APPLICANT: TRACI H. ROPP
  APPLICANT: LEIF P. SOLHEIM
  TITLE OF INVENTION: MUTANT ALPHABET
  NUMBER OF SEQUENCES: 40
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Genencor International
  STREET: 180 Kimball Way
  CITY: South San Francisco
  STATE: CA
  COUNTRY: USA
  ZIP: 94080
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/468700
  FILING DATE: 06-JUN-1995
  CLASSIFICATION: 252
ATTORNEY/AGENT INFORMATION:
  NAME: Stone, Christopher
  REGISTRATION NUMBER: 35,696
  REFERENCE/DOCKET NUMBER: GC27
TELECOMMUNICATION INFORMATION:

```

```
CC TELEPHONE: (415) 742-7555
CC TELEFAX: (415) 742-7217
CC INFORMATION FOR SEQ ID NO: 34:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 483 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
SQ SEQUENCE 483 AA; 55211 MW; 1221290 CN;

Query Match 68.0%; Score 2632; DB 1; Length 483;
Best Local Similarity 68.9%; Pred. No. 1.80e-219;
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;

Db 4 NGTLMQYFEWMPNDGQHWKRLQNDLSAYLAHGIHTAVWIPPAKGTSDQADVGAYDLYD 63
QY 37 NGTMMQYFEWHLDPNDGNHNRRLRDDAANLKSIGITAVWIPPAKGTSDQADVGAYDLYD 96
Db 64 LGFHHKGTVTKYGTGELQSAIKSLHSRDNVYGVVNNHKGADATEDVTAVEVDPA 123
QY 97 LGFHNKGTVTKYGTGELQSAIKSLHSRDNVYGVVNNHKGADATEDVTAVEVDPA 156
Db 124 DNRNVSIGELHKAHTHFFPGRGSTYSDFKWHWYHFDGTDWDSRKL-NRIYKFG--K 180
QY 157 NRNQEISGEYTIKAWTHFFPGRGSTYSDFKWHWYHFDGTDWDSRKL-NRIYKFG--K 216
Db 181 AWDWEVSNENGNIDYLYADIDYDHPDVAEIKRWGTWYANLQDGLDGFRLDAVKHIFSF 240
QY 217 AWDWEVDIENGNDYLYADIDYDHPDVAEIKRWGTWYANLQDGLDGFRLDAVKHIFSF 276
Db 241 LRQVNVHREKTKGEMTVAEYVNDLGALENYLNKTNFNSHVSFVPLHYQFHAASSTGG 300
QY 277 TRDLWTHVRNTTGKPMFAEFAEFKNDLAAENILNKTWNHSHVSFVPLHYQFHAASSTGG 336
Db 301 GYDMRLLNGTVSKHPLKSVTFVDNHDTPQGQSLSTVQWTFKPLAYAFILTRREGYPQ 360
QY 337 YFDMRNLNGSVVQKHPHIAVTFVDNHDTPQGQSLSTVQWTFKPLAYAFILTRREGYPQ 396
Db 361 VFYGDYIYGI--PTH-GVPSMKSIDPLQARQYAYGTQHDYFDHDIIGWTRGDSHP 420
QY 397 VFYGDYIYGI--PTH-GVPSMKSIDPLQARQYAYGTQHDYFDHDIIGWTRGDSHP 453
Db 421 NSGLAALITDGGPGAKRMVYGRONAGTWHDTGCRSEPVVINSSEGWFHNGGVSIIY 480
QY 454 NSGLATIMSDGPGGNKMWYVGHKAGQVWRDITGNRSGTGTINADGWNFTVNGGAVSVW 513
Db 481 VQR 483
QY 514 VKQ 516

RESULT 12
ID PCT-US95-10426-37 STANDARD; PRT; 487 AA.
XX
AC xxxxxx
XX
DE
DT
XX
DE
DE
XX
CC Sequence 37, Application PC/TUS9510426
CC GENERAL INFORMATION:
CC APPLICANT: GENENCOR INTERNATIONAL, INC.
CC TITLE OF INVENTION: An Improved Cleaning Composition
CC NUMBER OF SEQUENCES: 68
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genencor International
CC STREET: 180 Kimball Way
CC CITY: South San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94080
```

```
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/10426
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: STONE, Christopher L.
CC REGISTRATION NUMBER: 33,401
CC REFERENCE/DOCKET NUMBER: GC220-3
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 742-7536
CC TELEFAX: (415) 742-7217
CC INFORMATION FOR SEQ ID NO: 37:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 487 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 487 AA; 55495 MW; 1241380 CN;

Query Match 68.0%; Score 2632; DB 2; Length 487;
Best Local Similarity 68.9%; Pred. No. 1.80e-219;
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;

Db 8 NGTLMQYFEWMPNDGQHWKRLQNDLSAYLAHGIHTAVWIPPAKGTSDQADVGAYDLYD 67
QY 37 NGTMMQYFEWHLDPNDGNHNRRLRDDAANLKSIGITAVWIPPAKGTSDQADVGAYDLYD 96
Db 68 LGFHHKGTVTKYGTGELQSAIKSLHSRDNVYGVVNNHKGADATEDVTAVEVDPA 127
QY 97 LGFHNKGTVTKYGTGELQSAIKSLHSRDNVYGVVNNHKGADATEDVTAVEVDPA 156
Db 128 DNRNVSIGELHKAHTHFFPGRGSTYSDFKWHWYHFDGTDWDSRKL-NRIYKFG--K 184
QY 157 NRNQEISGEYTIKAWTHFFPGRGSTYSDFKWHWYHFDGTDWDSRKL-NRIYKFG--K 216
Db 185 AWDWEVSNENGNIDYLYADIDYDHPDVAEIKRWGTWYANLQDGLDGFRLDAVKHIFSF 244
QY 217 AWDWEVDIENGNDYLYADIDYDHPDVAEIKRWGTWYANLQDGLDGFRLDAVKHIFSF 276
Db 245 LRQVNVHREKTKGEMTVAEYVNDLGALENYLNKTNFNSHVSFVPLHYQFHAASSTGG 304
QY 277 TRDLWTHVRNTTGKPMFAEFAEFKNDLAAENILNKTWNHSHVSFVPLHYQFHAASSTGG 336
Db 305 GYDMRLLNGTVSKHPLKSVTFVDNHDTPQGQSLSTVQWTFKPLAYAFILTRREGYPQ 364
QY 337 YFDMRNLNGSVVQKHPHIAVTFVDNHDTPQGQSLSTVQWTFKPLAYAFILTRREGYPQ 396
Db 365 VFYGDYIYGI--PTH-GVPSMKSIDPLQARQYAYGTQHDYFDHDIIGWTRGDSHP 424
QY 397 VFYGDYIYGI--PTH-GVPSMKSIDPLQARQYAYGTQHDYFDHDIIGWTRGDSHP 453
Db 425 NSGLAALITDGGPGAKRMVYGRONAGTWHDTGCRSEPVVINSSEGWFHNGGVSIIY 484
QY 454 NSGLATIMSDGPGGNKMWYVGHKAGQVWRDITGNRSGTGTINADGWNFTVNGGAVSVW 513
Db 485 VQR 487
QY 514 VKQ 516

RESULT 13
ID PCT-US94-01553A-37 STANDARD; PRT; 487 AA.
XX
AC xxxxxx
XX
DE
DT
XX
```

DE Sequence 37, Application PC/TUS9401553A  
XX Sequence 37, Application PC/TUS9401553A  
CC GENERAL INFORMATION:  
CC APPLICANT: GENENCOR INTERNATIONAL, INC.  
CC TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase  
CC NUMBER OF SEQUENCES: 68  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Genencor International, Inc.  
CC STREET: 180 Kimball Way  
CC CITY: South San Francisco  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 94080  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US94/01553A  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Horn, Margaret A.  
CC REGISTRATION NUMBER: 33,401  
CC REFERENCE/DOCKET NUMBER: GC220-2  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 742-7536  
CC TELEFAX: (415) 742-7217  
CC INFORMATION FOR SEQ ID NO: 37:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 487 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC ORGANISM: Bacillus licheniformis  
CC SEQUENCE 487 AA; 55495 MW; 1241380 CN;

Query Match 68.0%; Score 2632; DB 2; Length 487;  
Best Local Similarity 68.9%; Pred. No. 1,80e-219;  
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;  
Db 8 NGLTMOYFEWYMPNDGQHWKRLQND SAYLA EHGITAVWIPPA YKGT SQADVGYGAYDLYD 67  
QY 37 NGTMOYFEWHLPNDCGNHNRLLRDDAANLKSXGITAVWIPPAWKGT SQNDVGYGAYDLYD 96  
Db 68 LGEFHQGTVRTKYGTGKELQSAIKSLHSRDINVTGVDVINHKGGADATEDVTAVEVDP 127  
QY 97 LGEFNQGTVRTKYGTGKELQSAIKSLHSRDINVTGVDVINHKGGADATEDVTAVEVDP 156  
Db 128 DRNRVISEGHLKAWTHFHPGSGTYSDFKWHYHFDGTDWDESRKL-NRIYKFGQ--K 184  
QY 157 NRNQEISGEYITEATWKFPDPPGRGNTHSNFKRWYHFDGTDWQSQQLQNKIYKFRGTGK 216  
Db 185 AWDWEYSNENGYDLYMAYADIDYDHPDVA AEIKRWGTWYANELQLDGFRLDAVKHKFSF 244  
QY 217 AWDWEVDIENGNYDLYMAYADIDMDHPEVINELRNNGVWYTNLTLDGFRIDAVKHKYSY 276  
Db 245 LRDNVNHVREKTKGEMFTVAEYWNQDLGALENYLNKTNTNHSVDFVPLHYQFHAASTQGG 304  
QY 277 TRDWTHTVNTTGPMPFAVEAFWKNDLAAIENYLNKTSWNHSHVDFVPLHYLNASNSGG 336  
Db 305 GYDMRKLNGTVVSKHPLKAVTFVDNHDTPQGSLSTVQTWFKPLAYAFILTRREGSYPQ 364  
QY 337 YFDMRNILGTVVQKHPITHAVTFVDNHDTPQGSLESTVQWTFKPLAYAFILTRREGSYPQ 396  
Db 365 VFYDGMYGTGKDSQREIPALKHKIEPILKARKQYAYGAQHDFDHDHDI VGTWREGDSSVA 424  
QY 397 VFYGDYGI--PTH-GVPSMKSIDPLQARQTYAGTQHDYFDHHDII GTWREGDSSHP 453  
Db 425 NSGLATITMSDGP GGNKMWYVKHKGAGQVVRDITGNRSCTVTINADGWNFTVNGGAVSVW 513

QY 454 NSGLATITMSDGP GGNKMWYVKHKGAGQVVRDITGNRSCTVTINADGWNFTVNGGAVSVW 513  
Db 485 VQR 487  
QY 514 VKQ 516  
RESULT 14  
ID US-08-645-971-2 STANDARD; PRT; 483 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE Sequence 2, Application US/08645971  
XX  
CC Sequence 2, Application US/08645971  
CC Patent No. 5763385  
CC GENERAL INFORMATION:  
CC APPLICANT:  
CC TITLE OF INVENTION: Modified Alpha-Amylases Having Altered  
CC FILING DATE:  
CC NUMBER OF SEQUENCES: 5  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/645,971  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 483 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC ORIGINAL SOURCE:  
CC ORGANISM: Bacillus licheniformis  
CC SEQUENCE 483 AA; 55268 MW; 1217764 CN;

Query Match 67.9%; Score 2629; DB 1; Length 483;  
Best Local Similarity 68.9%; Pred. No. 3.34e-219;  
Matches 333; Conservative 77; Mismatches 67; Indels 6; Gaps 4;  
Db 4 NGLTMOYFEWYMPNDGQHWKRLQND SAYLA EHGITAVWIPPA YKGT SQADVGYGAYDLYD 63  
QY 37 NGTMOYFEWHLPNDCGNHNRLLRDDAANLKSXGITAVWIPPAWKGT SQNDVGYGAYDLYD 96  
Db 64 LGEFHQGTVRTKYGTGKELQSAIKSLHSRDINVTGVDVINHKGGADATEDVTAVEVDP 123  
QY 97 LGEFNQGTVRTKYGTGKELQSAIKSLHSRDINVTGVDVINHKGGADATEDVTAVEVDP 156  
Db 124 DRNRVISEGHLKAWTHFHPGSGTYSDFKWHYHFDGTDWDESRKL-NRIYKFGQ--K 180  
QY 157 NRNQEISGEYITEATWKFPDPPGRGNTHSNFKRWYHFDGTDWQSQQLQNKIYKFRGTGK 216  
Db 181 AWDWEYSNENGYDLYMAYADIDYDHPDVA AEIKRWGTWYANELQLDGFRLDAVKHKFSF 240  
QY 217 AWDWEVDIENGNYDLYMAYADIDMDHPEVINELRNNGVWYTNLTLDGFRIDAVKHKYSY 276  
Db 241 LRDNVNHVREKTKGEMFTVAEYWNQDLGALENYLNKTNTNHSVDFVPLHYQFHAASTQGG 300  
QY 277 TRDWTHTVNTTGPMPFAVEAFWKNDLAAIENYLNKTSWNHSHVDFVPLHYLNASNSGG 336  
Db 301 GYDMRKLNGTVVSKHPLKAVTFVDNHDTPQGSLSTVQWTFKPLAYAFILTRREGSYPQ 360  
QY 337 YFDMRNILGTVVQKHPITHAVTFVDNHDTPQGSLESTVQWTFKPLAYAFILTRREGSYPQ 396  
Db 361 VFYDGMYGTGKDSQREIPALKHKIEPILKARKQYAYGAQHDFDHDHDI VGTWREGDSSVA 420  
QY 397 VFYGDYGI--PTH-GVPSMKSIDPLQARQTYAGTQHDYFDHHDII GTWREGDSSHP 453





\*\*\*\*\*  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPerch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Sep 16 10:26:09 1998; MasPar time 78.68 Seconds  
 1052.890 Million cell updates/sec

Tabular output not generated.

Title: >US-08-952-741-1  
 Description: (1-1776) from US08952741.seq  
 Perfect Score: 1776  
 N.A. Sequence: 1 atataaatttgaaatgaac.....ccaataaattggaagctt 1776  
 Comp: tatatttaacttacttgt.....gggtattattaaacttcgaa

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: n-issued  
 1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 8.900; Variance 4.431; scale 2.009

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Query Match % | Length | DB ID | Description                         | Pred. No. |
|------------|---------------|--------|-------|-------------------------------------|-----------|
| 1          | 376           | 21.2   | 5677  | 1 US-07-623- Sequence 4, Applicatio | 2.72e-289 |
| 2          | 374           | 21.1   | 1777  | 1 US-08-146- Sequence 33, Applicati | 1.38e-287 |
| 3          | 374           | 21.1   | 1968  | 2 PCT-US94-0 Sequence 31, Applicati | 1.38e-287 |
| 4          | 374           | 21.1   | 1968  | 2 PCT-US95-1 Sequence 31, Applicati | 1.38e-287 |
| 5          | 374           | 21.1   | 2149  | 1 US-07-623- Sequence 2, Applicatio | 1.38e-287 |
| 6          | 158           | 8.9    | 2182  | 3 5171673-5 Patent No. 5171673.     | 1.83e-105 |
| 7          | 98            | 5.5    | 7218  | 1 US-08-232- Sequence 14, Applicati | 1.31e-56  |
| 8          | 70            | 3.9    | 600   | 3 5171673-3 Patent No. 5171673.     | 1.24e-34  |
| 9          | 39            | 2.2    | 77    | 1 US-07-696- Sequence 8, Applicatio | 7.90e-12  |
| 10         | 39            | 2.2    | 481   | 3 5171673-1 Patent No. 5171673.     | 7.90e-12  |
| 11         | 32            | 1.8    | 105   | 1 US-07-865- Sequence 13, Applicati | 3.80e-07  |

|    |    |     |      |                                     |          |
|----|----|-----|------|-------------------------------------|----------|
| 12 | 30 | 1.7 | 7218 | 1 US-08-232- Sequence 14, Applicati | 7.26e-06 |
| 13 | 29 | 1.6 | 215  | 1 US-08-238- Sequence 5, Applicatio | 3.09e-05 |
| 14 | 26 | 1.5 | 215  | 1 US-08-238- Sequence 3, Applicatio | 2.12e-03 |
| 15 | 24 | 1.4 | 34   | 2 PCT-US95-1 Sequence 3, Applicatio | 3.17e-02 |
| 16 | 24 | 1.4 | 34   | 2 PCT-US94-0 Sequence 3, Applicatio | 3.17e-02 |
| 17 | 24 | 1.4 | 41   | 2 PCT-US95-1 Sequence 14, Applicati | 3.17e-02 |
| 18 | 24 | 1.4 | 41   | 2 PCT-US94-0 Sequence 14, Applicati | 3.17e-02 |
| 19 | 24 | 1.4 | 74   | 2 PCT-US95-1 Sequence 94, Applicati | 3.17e-02 |
| 20 | 24 | 1.4 | 81   | 2 PCT-US95-1 Sequence 92, Applicati | 3.17e-02 |
| 21 | 25 | 1.3 | 242  | 1 US-08-273- Sequence 1, Applicatio | 8.30e-03 |
| 22 | 23 | 1.3 | 74   | 2 PCT-US95-1 Sequence 100, Applicat | 1.18e-01 |
| 23 | 23 | 1.3 | 75   | 2 PCT-US95-1 Sequence 99, Applicati | 1.18e-01 |
| 24 | 23 | 1.3 | 81   | 2 PCT-US95-1 Sequence 98, Applicati | 1.18e-01 |
| 25 | 23 | 1.3 | 82   | 2 PCT-US95-1 Sequence 97, Applicati | 1.18e-01 |
| 26 | 22 | 1.2 | 41   | 2 PCT-US94-0 Sequence 17, Applicati | 4.24e-01 |
| 27 | 22 | 1.2 | 41   | 2 PCT-US94-0 Sequence 16, Applicati | 4.24e-01 |
| 28 | 22 | 1.2 | 41   | 2 PCT-US95-1 Sequence 29, Applicati | 4.24e-01 |
| 29 | 22 | 1.2 | 41   | 2 PCT-US94-0 Sequence 30, Applicati | 4.24e-01 |
| 30 | 22 | 1.2 | 41   | 2 PCT-US95-1 Sequence 16, Applicati | 4.24e-01 |
| 31 | 22 | 1.2 | 41   | 2 PCT-US95-1 Sequence 19, Applicati | 4.24e-01 |
| 32 | 22 | 1.2 | 41   | 2 PCT-US94-0 Sequence 24, Applicati | 4.24e-01 |
| 33 | 22 | 1.2 | 41   | 2 PCT-US94-0 Sequence 15, Applicati | 4.24e-01 |
| 34 | 22 | 1.2 | 41   | 2 PCT-US95-1 Sequence 25, Applicati | 4.24e-01 |
| 35 | 22 | 1.2 | 41   | 2 PCT-US95-1 Sequence 24, Applicati | 4.24e-01 |
| 36 | 22 | 1.2 | 41   | 2 PCT-US94-0 Sequence 25, Applicati | 4.24e-01 |
| 37 | 22 | 1.2 | 41   | 2 PCT-US95-1 Sequence 15, Applicati | 4.24e-01 |
| 38 | 22 | 1.2 | 41   | 2 PCT-US94-0 Sequence 19, Applicati | 4.24e-01 |
| 39 | 22 | 1.2 | 41   | 2 PCT-US95-1 Sequence 30, Applicati | 4.24e-01 |
| 40 | 22 | 1.2 | 41   | 2 PCT-US95-1 Sequence 17, Applicati | 4.24e-01 |
| 41 | 22 | 1.2 | 41   | 2 PCT-US94-0 Sequence 29, Applicati | 4.24e-01 |
| 42 | 22 | 1.2 | 66   | 2 PCT-US95-1 Sequence 93, Applicati | 4.24e-01 |
| 43 | 21 | 1.2 | 74   | 2 PCT-US95-1 Sequence 100, Applicat | 1.48e+00 |
| 44 | 22 | 1.2 | 81   | 2 PCT-US95-1 Sequence 98, Applicati | 4.24e-01 |
| 45 | 22 | 1.2 | 82   | 2 PCT-US95-1 Sequence 97, Applicati | 4.24e-01 |

# ALIGNMENTS

RESULT 1  
 ID US-07-623-953-4 STANDARD; DNA; UNC; 5677 BP.  
 AC xxxxxx

DE Sequence 4, Application US/07623953  
 CC Sequence 4, Application US/07623953  
 CC Patent No. 5364782

CC GENERAL INFORMATION:

CC APPLICANT: Quax, Wilhelmus J  
 CC APPLICANT: Laroche, Yves  
 CC APPLICANT: Vollebregt, Adrianus W.H.  
 CC APPLICANT: Stanssens, Patrick  
 CC APPLICANT: Lauwereys, Marc

CC TITLE OF INVENTION: MUTANT MICROBIAL ALPHA-AMYLASES WITH  
 CC TITLE OF INVENTION: INCREASED THERMAL, ACID AND/OR ALKALINE STABILITY  
 CC NUMBER OF INVENTION: 7

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM  
 CC STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR  
 CC CITY: PALO ALTO  
 CC STATE: CA

CC COUNTRY: USA

CC ZIP: 94306

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

***This Page Blank (uspto)***



***This Page Blank (uspto)***

Sep 16 10:24

US-08-952-741-1 mi

5

RESULT 2  
ID US-08-146-422-33 STANDARD; DNA; UNC; 1777 BP.  
AC xxxxxx  
DT Sequence 33, Application US/08146422  
CC Sequence 33, Application US/08146422  
CC Patent No. 5543576  
CC GENERAL INFORMATION:  
CC APPLICANT: VAN OOLJEN, ALBERT J. J.  
CC APPLICANT: RIETVELD, KRILIN  
CC APPLICANT: HOEKEMA, ANDREAS  
CC APPLICANT: PEN, JAN  
CC APPLICANT: SIJMONS, PETER C.  
CC APPLICANT: VERHOERD, TEUNIS C.  
CC APPLICANT: QUAX, WILHEMUS J.  
CC TITLE OF INVENTION: PRODUCTION OF ENZYMES IN SEEDS AND THEIR  
CC TITLE OF INVENTION: USE  
CC NUMBER OF SEQUENCES: 33  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: MORRISON & FOERSTER  
CC STREET: 755 Page Mill Road  
CC CITY: Palo Alto  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94304-1018  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC FILING DATE: 02-NOV-1993  
CC CLASSIFICATION: 514  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: KENNEDY, BILL  
CC REGISTRATION NUMBER: 33,407  
CC REFERENCE/DOCKET NUMBER: 44615-20011.23  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 813-5600  
CC TELEFAX: (415) 494-0792  
CC TELEX: 706141  
CC INFORMATION FOR SEQ ID NO: 33:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1777 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
SQ SEQUENCE 1777 BP; 500 A; 366 G; 462 C; 449 T; 0 OTHER.

Query Match 21.1%; Score 374; DB 1; Length 1777;  
Best Local Similarity 66.1%; Pred. No. 1.38e-287;  
Matches 951; Conservative 0; Mismatches 469; Indels 18; Gaps 9;  
DB 106 AATGGACGCTGATGACGATTTTGAATGCTGACATGCCCAATGACGGGCAACATTTGAAG 165  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
QY 253 aatggaccatgatgcaglattttgatggcatttgccaaatgacgggaaccactggagac 312  
DB 166 CGTTTGCAAAACGACTCGGCATATTTGGCTGAACACGGCTATTACTGCGCTCTGGATTCCC 225  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 313 aggttaagagatgacgcagctaaacttaagagtaagggttaaccgtgtttgattccet 372  
DB 226 CCGGATATTAAGGAACGACGACGGGATCTGGGCTACGGCTTTACGACCTTTATGAT 285

Sep 16 10:24

US-08-952-741-1 mi

6

QY 373 cctgcatggaaggggacttcgcaaaatgatgtgggtatggtgectatgatttgtacgat 432  
DB 286 TTAGGGAGTTTCATCAAAAAGGCGCGTTTCGCAAAAGTAGCGCACAAGAGAGAGCTG 345  
||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
QY 433 cttggtagtttaaccaaagggaaccgtccgtacaaaatattggcacaaggagtcagttg 492  
DB 346 CAATCTGCGATCAAAAAGTCTTCATTCGCCGACGACATTAAAGTTTACGGGGATGTGGTCA 405  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 493 caaggtgcgtagcatctttgaaaaataaacgggattccaagttttatggggatgctgtagt 552  
DB 406 AACCAAAAGGCGCGCTGATGCGACCGAAGATGTAACCGCGTTGAAGTCGATCCCGCT 465  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 553 aatcataaaggtggagcagcgggacagagatggaatgcggtggaagtgaaccgaagc 612  
DB 466 CACGCAACCGCGTAATTTTCAGGAGAACCTTAATTAAGCCCTGGACACATTTTCATTTT 525  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 613 aaccgaaccaagaataatcaggtgaataccattgaagcatggaacaaatttgatttc 672  
DB 526 CCGGGCGCGGACGACATACAGCGATTTTAAATGCGATTGGTACCAATTTTGACGGAAC 585  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 673 cctggaaggaggaaataaccattccaactttaaatggcggtggtatcatcttttgatggaca 732  
DB 586 CATTGGGACGAGTCCGAAAAGT---GAACCGCATCTATAAGTTTC-A--A--GAAAG 636  
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
QY 733 gattgggactcagtcagctcagtcagcaaaaaataataaattcagaggaccggaag 792  
DB 637 CTTTGGGATGGGAACTTTTCAATGAAAACGCACTATGATTATTTGATGATCGGAC 696  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 793 gcatgggactgggaagttagatagagaacggcaactatgattacccttatgtatgcagac 852  
DB 697 ATCGATTATGACCATCCTGATGTCGACGAGAAATTAAGAGATGGGCACTTGGTATGCC 756  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 853 attgatagatcatcagaagtaatacaatgaacttagaattgggaattgggtatata 912  
DB 757 AATGAATGCAATTGGACGGTTCCGCTCTGATGCTGTCAACACATTAATTTCTTTT 816  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 913 aatacacttaactagatggattagaatcgatgctggaacataattaataacagctat 972  
DB 817 TCGGGGATTTGGTTAATCATGTCAGGAAAAAAGCGGCAAGCAATGTTTACGTA 876  
||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 973 acgagagattggtaacacatgctgtaacaccacaggtaaaccaatgtttgcagttgca 1032  
DB 877 GAATATTGCGAAGTACATTGGGCGCTGGAAACTATTTCACAAAACAAATTTTAAT 936  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 1033 gaattttggaaaaatgacctgtgtgcaatcgaaaactatttaataaacaagttggaa 1092  
DB 937 CATTGAGTGTGACGTCGCGCTTCATTATCATGTTCCATGCTGTCATGACACAGGAGGC 996  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 1093 cactccgtgttcagttctctcttcatttaattgtacaatgcatcctaataatggtggc 1152  
DB 997 GGCTATGATGAGGAAATGCTGAACGGTAGCGTCTGTTTCAACGATCCGTTGAATCG 1056  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 1153 tatttgatagaaaatattttaaatggttctgtcgtaaaaaaacccctatacatgca 1212  
DB 1057 GTTACATTTGTGATACCATGATACACACGCGGGCAATCGCTTGACTCGACTGTCAA 1116  
||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 1213 gtcaattgttgataaccatgactctcagccaggagagaagcattggaatcctttgttcaa 1272  
DB 1117 ACATGTTTAAGCCGCTTGTACCGCTTTTATCTCACAAGGGAATCTGGATACCTCAG 1176  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY 1273 cgtggttcaaacaccactggcatatgcattgattcttgacaaggagggaaggtaccctcc 1332  
DB 1177 GTTTTCTACGGGGATATGTACGGGACGAAAGGAGACTCCCGACGGCAAAATTCCTGCTTG 1236

***This Page Blank (uspto)***

QY 1333 gattttacggtgattactacggtat---acca-actcat-g-gtg---tcccttcgatg 1383  
 Db 1237 AAACACAAATTCACCGATCTTAAACGAGAAACAGTATGCTACGAGACAGCAT 1296  
 QY 1384 aaatcaaatgtaccactctgcagcgctcaacgctatgctcactacggaaccaat 1443  
 Db 1297 GATTATTTCACCAACATGACATGCTCGCTGACAGGAGGACGACAGCTCGGTGCA 1356  
 QY 1444 gattatttgatcatcatgatattatcggtgagcgagagaaggagacagctccaccoca 1503  
 Db 1357 AATTGAGTTTGGCGGATTAATAACAGACGACCGCTGGGGCAAGCAATGTATGTC 1416  
 QY 1504 aattcaggatgcaactattatgctcagatggcgagggggaataaatgagatgatgctc 1563  
 Db 1417 GCGCGCAAAACCGCGGTGACATGGCATGACATTACCGGAAACCGTTTCGAGCGGTT 1476  
 QY 1564 ggggaacataaagctggcgaagtatgagagatatcacgggaataataggtctggtaccgtc 1623  
 Db 1477 GTCATCAATTCGGAAGCGTGGGAGAGTTTCAGCTAAACGCGGGTTCGTTCAATTT 1534  
 QY 1624 accattaatgcagatggttgggggaatttcacgtgaaacggagggcaggttcggtt 1681

RESULT 3  
 ID PCT-US94-01553A-31 STANDARD; DNA; UNC; 1968 BP.  
 AC xxxxxx  
 DT  
 DE Sequence 31, Application PC/TUS9401553A  
 CC Sequence 31, Application PC/TUS9401553A  
 CC GENERAL INFORMATION:  
 CC APPLICANT: GENECOR INTERNATIONAL, INC.  
 CC TITLE OF INVENTION: Oxidatively Stable Alpha-Amylase  
 CC NUMBER OF SEQUENCES: 68  
 CC  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Genecor International, Inc.  
 CC STREET: 180 Kimball Way  
 CC CITY: South San Francisco  
 CC STATE: CA  
 CC COUNTRY: USA  
 CC ZIP: 94080  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent In Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/US94/01553A  
 CC FILING DATE:  
 CC CLASSIFICATION:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Horn, Margaret A.  
 CC REGISTRATION NUMBER: 33,401  
 CC REFERENCE/DOCKET NUMBER: GC220-2  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (415) 742-7536  
 CC TELEFAX: (415) 742-7217  
 CC INFORMATION FOR SEQ ID NO: 31:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 1968 base pairs  
 CC TYPE: nucleic acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: DNA (genomic)

QY SEQUENCE 1968 BP; 573 A; 388 C; 509 G; 498 T; 0 OTHER.  
 Query Match 21.1%; Score 374; DB 2; Length 1968;  
 Best Local Similarity 66.1%; Pred. No. 1.38e-287;  
 Matches 951; Conservative 0; Mismatches 469; Indels 18; Gaps 9;  
 Db 260 AATGGGACCGCTGATGCGATTTTGAATGTACATGCCAATACGCGCCAAATTTGGAAG 319  
 QY 253 aatgggacatgatgcagtcatttgaatggcatttgcgaatgacggggaaccaactggaac 312  
 Db 320 GCTTTCAAAACGACTCGGCATATTTGGCTGAACACGGTATTACTGCGCTCGGATTCGC 379  
 QY 313 aggttacgagatgacgcagtaacttaagagtaagaggattaccgctgttggtattcct 372  
 Db 380 CCGGCTATATAGGACGAGCGAAGCGGATGTGGCTAGCGTCTAGCACCTTTATGAT 439  
 QY 373 cctgcataagaggacttcgcaaaatgatgtgggtatggtgcctatgattgtacgat 432  
 Db 440 TTAGGGAGTTTCATCAAAAAGGACGCTTCGGCAAAAGTACGGCAAAAAGGAGAGCTG 499  
 QY 433 ctgtgtgagtttaacaaaagggaacgctccgacaaaatatggcacaaggagtcagttg 492  
 Db 500 CAATCTCCGATCAAAAAGTCTTCATTCGCCGACATTTAACTGTTACGGGGATGTGCTATC 559  
 QY 493 caaggtgccgtgacatcttctgaaaaataacgggattccaagtttatggggatgctgtag 552  
 Db 560 AACCACAAAAGCGCGCTGATGCGACCGAAGATGTAACCGCGGTTGAAGTCGATCCGCT 619  
 QY 553 aatcataaagdtggagcagcagggacagagatggttaatgcggtggaagtgaaccgaagc 612  
 Db 620 GACCGCAACCGCGTAAATTCAGGAGAACACCTTAATTAAGCCTGACACATTTTCAATTT 679  
 QY 613 aaccgaacccaagaatatcaggatgaatacacattgaacatggaacgaatttgatttc 672  
 Db 680 CCGGGCGCGGACGACCATACAGCGATTTTAAATGGCATTTGTGTACCATTTTCAGCAACC 739  
 QY 673 cctggaagagaaataacccttccaactttaaagcgctggtatcatttggatggaca 732  
 Db 740 GATTGGACGAGTCCCGAAAGCT---GAACCGCATCTATAAGTTTC-A---A---GGAAG 790  
 QY 733 gattggagatcagtcagtcagcttcagacaaaataataaattcagaggtaccggaag 792  
 Db 791 GCTTGGGATTGGGAAGTTTCAATCAAAAACGCGCACTATGATTATTTCATGTCGCCAC 850  
 QY 793 gcatgggactgggaagtagatatagagaacggcaactatgattaccctatgtatgcagac 852  
 Db 851 ATCGATTATGACCATCTGATGCTCCGACGAGAAATTAAGAGATGGGCGACTTGGTATGCC 910  
 QY 853 attgatattgatcatccagaagtaaatcaatgaacttagaaattgggagttgggtatata 912  
 Db 911 AATCAACTGCAATTGACCGGTTTCGCTTGTATGCTCTCAACACATTAATAATTTCTTTT 970  
 QY 913 aatacacttaactagatgatttgaatcgatgctgtgaacataataataacacagctat 972  
 Db 971 TTGCGGGATTGGGTTATCATGTCAGGGAANAACGGGGAAGAAATGTTTACGTAGCT 1030  
 QY 973 acgagagattggtcaacacatgctgcgaacacacaggttaacccaatgtttgcagttgca 1032  
 Db 1031 GAATATTGGCAGAAATGACTTGGCGCGCTGGAATAACTATTGGAACAAAACAAATTTAAT 1090  
 QY 1033 gaattttgaaaaaatgacctgtgcaatcgaaaactatttaataaacaagaagtggaaat 1092  
 Db 1091 CATTCAAGTCTTGGCGGCTTCATTATCAGTTTCCATCTGCTATCCACACAGGGAGGC 1150



***This Page Blank (uspto)***



***This Page Blank (uspto)***



***This Page Blank (uspto)***

Db 718 CAATCTGCATCAAAAGTCTTCAATCCCGCAGCAATTAAGCTTTACGGGGATGTGTCATC 777  
 Qy 493 caagggtgcggtgacatcttggaaataaaccggattcaagtttgaattggggtatgctggtg 552  
 Db 778 AACCAAAAGCGCGGCTGATCGCCGCAAGATGATTAACCGCGGTTGAGTCTGATCCCGCT 837  
 Qy 553 aatcataaagtgagacagcgagacagagatggttaaatgcggtggaagtgaaccgaagc 612  
 Db 838 GACCGCAACCGCTTAATTTGAGGAGAACACCTAATTAAGCCCTGGACACATTTTCATTTT 897  
 Qy 613 aaccgaaccaagaataatcaggtgaatacaccattgaagcatgagcgaatttgatttc 672  
 Db 898 CCGGGCGCGGAGACATACAGCGAATTTAAATGGCATTTGGCATTTTACGGAAC 957  
 Qy 673 cctggaagagaataaccattccaacttaaatgagctggtatcaatttgatggagaca 732  
 Db 958 GATTGGAGCAGTCCGAAAGCT---GAACCGCATCTATAAGTTTC-A---A---GGAAG 1008  
 Qy 733 gattggatcagtcacgtccagtcacgaacaaataataaattcagaggtaccgggaag 792  
 Db 1009 GCTTGGGATTGGGAGTTCGAATGAAGGCAACTATGATTAATTTGATGATGCGGAC 1068  
 Qy 793 gcatggactggaagatagatagagaacggcaactatgattaccttatgtatgcagac 852  
 Db 1069 ATCGATTATGACATCCGTGATGCGCAGCAGAAATTAAGAGATGGGCACTGGTATGCC 1128  
 Qy 853 attgataggatcatccagagaataatcaatgaacttagaaattgggagtttggtataca 912  
 Db 1129 AATGAACATGCAATGGAGGTTCCGCTTGTGATGCTGCAACACATTAATTTCTTTT 1188  
 Qy 913 aatacactaatctagatggatttagaatcgatgctggaacataataataacagctat 972  
 Db 1189 TTCGGGATTTGGGTAAATGTCAGGGAATAAAGCGGAGGAATGTTTACGGTAGCT 1248  
 Qy 973 acgagagattggtcaacacatgctgaacacacaggttaacacatggttgaggtgca 1032  
 Db 1249 GAATATTGGCAAAATGACATTTGGGCGGCTGGAAACTATTTTGAACAAACAAATTTAAT 1308  
 Qy 1033 gaatttggaaatgacctgctgcaatcgaaactatttaataaacaagaattggaa 1092  
 Db 1309 CATTGATGTTTGAATGCTGCGCTTCAATATGATGCTGATGCTGATGCTGCAACAGGAGG 1368  
 Qy 1093 cactcgtgttcgattgctcctcttcaataaattgtacaatgcatctaatagtggtgc 1152  
 Db 1369 GGCTATGATAGGAAATTCGTGACGCTAGCGTCTTCCAGACATCGGTTGAATCG 1428  
 Qy 1153 tattttgatagaaatatttaaatggtctgctgacaaaacacctatacatgca 1212  
 Db 1429 GTTACATTTGTCATTAACATGATACAGCGCGGCAATCGTTGATGCTGATGCTGCA 1488  
 Qy 1213 gtcaatttggataaaccatgactctcagccagagagaagcattggaactcttggttcaa 1272  
 Db 1489 ACATGTTTAAAGCGGTTGCTTACGCTTTTATCTCACAAGGGAATCTGATACCCCTCAG 1548  
 Qy 1273 tctgggtcaaacacctggcatatgattgattctgacaagggagaggttacccttcc 1332  
 Db 1549 GTTTTCTACGGGATGATGACGGAGCAAGGACATCCCGAGCGAATTCCTGCGCTTG 1608  
 Qy 1333 gtattttcaggtgattactacggtat---acca-actcat-g-gtg---ttccttcgatg 1383  
 Db 1609 AAACACAAATTCACCGATCTTAAGAGGAGAAACAGTATGCTGCTAGCAGCACACAT 1668  
 Qy 1384 aaatctaaatgatccactctgcaggcagctcaaacgtatgctacaggaaccaacat 1443

Db 1669 GATTATTTCACCATGACATCTCGCTGCAGCAGGAAGGCGCAGCTCGGTTGA 1728  
 Qy 1444 gattatttgcacatcatgattatcgctgagcagagaagaagggagacagctcccacca 1503  
 Db 1729 AATTGAGTTTCGGGCGCATTAATACACAGCGGACCGCGTGGGCAAGCGAATGTATGTC 1788  
 Qy 1504 aattcaggacttgcaactattatgtccgatggccagggggttaataaatggatgtatgct 1563  
 Db 1789 GCGCGGCAAAAGCGGCTGAGCATGACATGACATACCGGAACCGTTCGGAGCGGTT 1848  
 Qy 1564 ggaatacataaagctggccaagtatgagagatatcacccggaataagctctgtaccgctc 1623  
 Db 1849 GTCAATCAATTCGAAGGCTGGGAGAGTTTCACGTAAACGGGGTTCGTTTCAATTT 1906  
 Qy 1624 accattaatgcagatggttgggggaatttcaactgtaaacgaggggcagtttcggttt 1681  
 RESULT 6  
 ID 5171673-5 STANDARD; DNA; UNC; 2364 BP.  
 AC xxxxxx  
 DT 01-JAN-1900  
 DE Patent No. 5171673.  
 CC Patent No. 5171673  
 CC APPLICANT: SLOMA, ALAN; HANNETT, NANCY M.; STEPHENS, M.A.  
 CC RUDOLPH, CATHY F.; RUFO JR., GERALD A.; PERO, JANICE  
 CC TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS DNA USING  
 CC THE BACILLUS COAGULANS AMYLASE GENE  
 CC NUMBER OF SEQUENCES: 10  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/219,599  
 CC FILING DATE: 18-JUL-1988  
 CC SEQ ID NO:5:  
 CC LENGTH: 2182  
 CC Sequence 2364 BP; 611 A; 482 C; 528 G; 561 T; 182 other;  
 SQ  
 Query Match 8.9%; Score 158; DB 3; Length 2182;  
 Best Local Similarity 58.7%; Pred. No. 1.83e-105;  
 Matches 651; Conservative 0; Mismatches 451; Indels 7; Gaps 7;  
 Db 376 GAATCATPACATCATGCGCTTTTGAATGGAATACCCAGCAGACGCGCAATGGAA 435  
 Qy 252 gaatggaccatgcatgatttgaatggcatttgcacaaatgacgggaaccactggaa 311  
 Db 436 CCGGCTCAAGAAATGGCGCTGAATTAAGAAAGGCGGATTTGATGCGCTCTGGGTTCC 495  
 Qy 312 caggttacgagatgacgcagctaaacttaagagtaaggattaccgctgtttggattcc 371  
 Db 496 CCGGCTCAAAAAGCAGCTCAGCATGACATGGAATGTTACGGGCTGTACGACCATTTACA 555  
 Qy 372 tctcgatggaaggagcttcgcaaaatgatgttggtatggtgcctatgatttgaaga 431  
 Db 556 CCGCGGAGTTTACCAAGAAAGCGGCTCAGGACAAAGTAGTCGGGACAAACAGCAATT 615  
 Qy 432 tcttggtgagtttaaccaaagggaaccgtccgtacaaaataatggcacaaggagtcagt 491  
 Db 616 ACATGAAGCGATCAACCGCTGCCAGCAGCATATCCAGGTCTATATCGATGCTCAT 675  
 Qy 492 gcaaggtgcgctgacatctttgaaaaataaacgggatccaagtttatggggtatgctgat 551  
 Db 676 GAACCATTAAGCGCGCGGATGAACCCGAATCTTTTCCAAAGTGTGTGAGGTGCGACCGCAT 735  
 Qy 552 gaatcataaaggtggagcagcgagggacagagatggttaaatgcggtggaagtgaaccgaag 611  
 Db 736 GGACCGCAACAAAGAAATTTCCGAAACGGTTTGAATGAAGGCTGGACAAAGTTCAATTT 795

***This Page Blank (uspto)***



Sep 16 10:24

US-08-952-741-1 mi

15

```

||||| ||| ||||||| || | | |||||| |||||| ||| ||| |||
Qy 612 caaccgaacaaagaataatcagggaatacacacattgaagcatggacgaataattgattt 671
Db 796 TACAAACCGGAGGAAAGGCAATGTTCTGATTTCGTGGAATATACACCATTCAGTGGCGT 855
Qy 672 ccttgaagagggaataaccattccaaacttaaatggcgttggtatcatttggatggac 731
Db 856 CGATTACGA-CAACCG-GACAGCG-CGGAACGGCATT-TCCGTATTGTCGGGGAATA 911
Qy 732 agattggatcagtcacgtcagttcagaacaaataataaattcagaggtaccggaaa 791
Db 912 ACATTGCGATGAGCATGCGAACGCAATTCGAACTTCGATTATTCATGCTACGCGG 971
Qy 792 ggcattgggactggaaa-gtagatatagaaacggcaactatgattacccttatgtatgcag 850
Db 972 ACATTGATTAACACCATCCGATGTAAAGAAAGAAATGATCGAATGGGGAATATGGCTGG 1031
Qy 851 acattgatagatcatccagaagtaataatgaacttagaattgggaggtttggtata 910
Db 1032 CCGATACGCGGTTGCGACGGCTACGGCTCGATCCCATTAAGCATATCAACCATGACT 1091
Qy 911 caatacacttaacttagatggatttagaatcgtgctgtgaaacatatataacagct 970
Db 1092 TTATCCGCGACTTTCGCGCTGCTCTTAATGGAACACCGC-GGAGACCATTTTATTTTGTC 1150
Qy 971 atacgagaga-ttggctaacacatgtcgttaaccacaggtaaaccaatgtttgcagtt 1029
Db 1151 GCGGAGTTCTGGAATCCCGACTTGGACGGTGCAGAAATACCTCGACCATGTACAGTTT 1210
Qy 1030 gcagaattttggaaaatgacctgtcgtcaatcgaaactatttaataaacaagtgtg 1089
Db 1211 AAMATCGATTGTTGATGTTGCACTCCATTAATAATTCGATGAGCATGTAAAGAGG 1270
Qy 1090 aatcactcgtgtgagttcctctcattataattgtacaatgcactaatagttg 1149
Db 1271 CCGCGGTTGACCTCCGACGATTTTTCATGATACGCTGTTCAACGACCGCGTGAAT 1330
Qy 1150 ggcattttgatagagaataatttaattgtctgtcgtaacaaacacccatacat 1209
Db 1331 GCTGTACGTTGTCGATAACCATGATTACAGCCGACGAATCGCTGGAATCATGGTG 1390
Qy 1210 gcagtcacattttgtataacacatgactctcagccaggaggaagcattggaatccttgtt 1269
Db 1391 GACGATGTTTAAACGACGCGCTATGCTGTTGATTCTGCGGAAAGACGGTTATCGG 1450
Qy 1270 caatcgtgttcaaacactggtcatgattgattcgtgacaaggaggaaggttaacct 1329
Db 1451 TCGGCTTTTACGGGATGATGACGGCAT 1479
Qy 1330 tccgatttttaccgtgtattacaggtat 1358

```

RESULT 7

ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.

AC xxxxx

DT

DE Sequence 14, Application US/08232463

CC Sequence 14, Application US/08232463

CC Patent No. 5670367

CC GENERAL INFORMATION:

CC APPLICANT: DORNER, F.

CC APPLICANT: SCHEFLINGER, F.

CC APPLICANT: FALKNER, F. G.

CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

Sep 16 10:24

US-08-952-741-1 mi

16

```

CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232,463
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/935,313
CC FILING DATE:
CC APPLICATION NUMBER: EP 91 114 300.6
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 836-9300
CC TELEFAX: (703) 683-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: pTZgpt-Fls
CC SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

```

```

Query Match 5.5%; Score 98; DB 1; Length 7218;
Best Local Similarity 1.1%; Pred. No. 1.31e-56;
Matches 4; Conservative 227; Mismatches 133; Indels 0; Gaps 0;

```

```

Db 1075 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1134
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Cc 827 ttgcgtctctatctactctccatgccatgcttccgttacctctgaatttat 768
Db 1135 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1194
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Cc 767 attttgtcgaagtcgagtcgactgacccaactgtcccatcaaaatgataccagcgc 708
Db 1195 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1254
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Cc 707 catttaagtgggaatgggtattctctccagggaatacaaatcttcgcatcctca 648
Db 1255 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1314
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Cc 647 atgggtattcacctgatattctgttcggtgcttcggttcattccaccgcat 588
Db 1315 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1374
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Cc 587 accatctctgcccgtctgcccactttatgattcatcacgacatcccccaaaactga 528

```

***This Page Blank (uspto)***

Sep 16 10:24

US-08-952-741-1 mi

17

Db 1375 YYY 1434  
 Cp 527 atccgtatttttcaagatgacgcggcaacttgcaactgactcctctgtgcataattt 468  
 Db 1435 GTAC 1438  
 Cp 467 gtac 464

RESULT 8

ID 5171673-3 STANDARD; DNA; UNC; 650 BP.  
 AC xxxxxx  
 DT 01-JAN-1900  
 DE Patent No. 5171673.  
 CC Patent No. 5171673  
 CC APPLICANT: SLOMA, ALAN; HANNETT, NANCY M.; STEPHENS, M.A.  
 CC RUDOLPH, CATHY F.; RUFO JR., GERALD A.; PERO, JANICE  
 CC TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS DNA USING  
 CC THE BACILLUS COAGULANS AMYLASE GENE  
 CC NUMBER OF SEQUENCES: 10  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/219,599  
 CC FILING DATE: 18-JUL-1988  
 CC SEQ ID NO:3:  
 CC LENGTH: 600  
 CC Sequence 650 BP; 189 A; 124 C; 135 G; 152 T; 50 other;

Query Match 3.9%; Score 70; DB 3; Length 600;  
 Best Local Similarity 65.2%; Pred. No. 1.24e-34;  
 Matches 150; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Db 315 GAATCATACATCATGCGATTTTGAATGGAATACGCCAGCAGCGCCATTGGAA 374  
 Qy 252 gaatggaccatgatgcagattttgaatggcatttgcacaaatgacgggaaccactgaa 311  
 Db 375 CCGGCTGAAGAAATGGCGCCCTGAATTAAGAAAGCGGATTCATGCGCTCTGGCTCC 434  
 Qy 312 caggttacagatgacgcagctaaactaaagatgaaaggtattcagctgttgattcc 371  
 Db 435 GCGGTGACAAAGACAGTACAGATGACATGTTACGGGGTGTACCAACCATACGA 494  
 Qy 372 tctgcatggaagggaacttcgcaaatgatgttggtgatggtcctatgattgtacga 431  
 Db 495 CTCTGGGCGATTTCACCAAGAGCCCTGACGACAAAGTACGGGACAA 544  
 Qy 432 tcttggtgagtttaaccaaagggaacgcgtccgtacaaaatattggcaaa 481

RESULT 9

ID US-07-696-551B-8 STANDARD; DNA; UNC; 77 BP.  
 AC xxxxxx  
 DT  
 DE Sequence 8, Application US/07/696551B  
 CC Sequence 8, Application US/07/696551B  
 CC Patent No. 5232841  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Hashimoto, Tamotsu  
 CC APPLICANT: Tsujimura, Atsushi  
 CC APPLICANT: Ueda, Shigezo  
 CC TITLE OF INVENTION: Process for Preparing Peptide  
 CC NUMBER OF SEQUENCES: 12  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
 CC ADDRESSEE: Dunner

Sep 16 10:24

US-08-952-741-1 mi

18

CC STREET: 1300 I Street, N.W., Suite 700  
 CC CITY: Washington  
 CC STATE: D.C.  
 CC COUNTRY: USA  
 CC ZIP: 20005-3315  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: MS-DOS/PC-DOS  
 CC SOFTWARE: Patent in Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/696,551B  
 CC FILING DATE: 19910509  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 2-122166  
 CC FILING DATE: 11-MAY-1990  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 2-334575  
 CC FILING DATE: 30-NOV-1990  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Lawrence M. Lavin, Jr.  
 CC REGISTRATION NUMBER: 30,768  
 CC REFERENCE/DOCKET NUMBER: 2481-1070  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (202) 408-4000  
 CC TELEFAX: (202) 408-4400  
 CC INFORMATION FOR SEQ ID NO: 8:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 77 bases pairs  
 CC TYPE: NUCLEIC ACID  
 CC STRANDEDNESS: both  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: DNA (genomic)  
 CC SEQUENCE 77 BP; 22 A; 16 C; 21 G; 18 T; 0 OTHER.

Query Match 2.2%; Score 39; DB 1; Length 77;  
 Best Local Similarity 83.1%; Pred. No. 7.90e-12;  
 Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db 17 AATGGACCTGATGTCAGTATTGTAATGCTACATGCCAATGACGGCCAACTTGGAA 75  
 Qy 253 aatggaccatgatgcagtatttgaatggttcattgccaatgacgggaaccactggaa 311

RESULT 10

ID 5171673-1 STANDARD; DNA; UNC; 521 BP.  
 AC xxxxxx  
 DT 01-JAN-1900  
 DE Patent No. 5171673.  
 CC Patent No. 5171673  
 CC APPLICANT: SLOMA, ALAN; HANNETT, NANCY M.; STEPHENS, M.A.  
 CC RUDOLPH, CATHY F.; RUFO JR., GERALD A.; PERO, JANICE  
 CC TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS DNA USING  
 CC THE BACILLUS COAGULANS AMYLASE GENE  
 CC NUMBER OF SEQUENCES: 10  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/219,599  
 CC FILING DATE: 18-JUL-1988  
 CC SEQ ID NO:1:  
 CC LENGTH: 481  
 CC Sequence 521 BP; 145 A; 102 C; 96 G; 138 T; 40 other;  
 Query Match 2.2%; Score 39; DB 3; Length 481;

***This Page Blank (uspto)***

|                       |   |  |                                     |
|-----------------------|---|--|-------------------------------------|
| Best Local Similarity | 68.0%;  | Pred. No. 7,90e-12;  |                                     |
| Matches               | 71;   | Conservative   | 0; Mismatches 32; Indels 0; Gaps 0; |
| Db                    | 376   | GAATCATACATCATGCTGCTTTTGAATGCGAATACGCCAGACAGCGGCGCATTTGGAA | 435                                 |
| Qy                    | 252   | gaatggccatgatgcagtatattgaatgcttgcgaatgacgggaaccactggaa     | 311                                 |
| Db                    | 436   | CGCGCTCAAGAAATGCCGCCCTGAATTAAAGAAAGCGCGATT                 | 478                                 |
| Qy                    | 312   | caggttcagatgacgcagctaaacttaagagtaaggagatt                  | 354                                 |
| RESULT                | 11  |  |                                     |
| ID                    | US-07-865-662F-13   | STANDARD; DNA; UNC; 105 BP.                                |                                     |
| AC                    | xxxxxx  |  |                                     |
| DT                    |   |  |                                     |
| DE                    | Sequence 13, Application US/07865662F                                   |  |                                     |
| CC                    | Sequence 13, Application US/07865662F                                   |  |                                     |
| CC                    | Patent No. 5451670  |  |                                     |
| CC                    | GENERAL INFORMATION:  |  |                                     |
| CC                    | APPLICANT: Marcia M. Miller   |  |                                     |
| CC                    | TITLE OF INVENTION: Restriction Fragment Length                         |  |                                     |
| CC                    | TITLE OF INVENTION: Polymorphism Test For Haplotyping Domesticated Fowl |  |                                     |
| CC                    | NUMBER OF SEQUENCES: 14   |  |                                     |
| CC                    | CORRESPONDENCE ADDRESSES:   |  |                                     |
| CC                    | ADDRESSEE: City of Hope   |  |                                     |
| CC                    | STREET: 1500 East Duarte Road   |  |                                     |
| CC                    | CITY: Duarte  |  |                                     |
| CC                    | STATE: California   |  |                                     |
| CC                    | COUNTRY: United States of America                                       |  |                                     |
| CC                    | ZIP: 91010-0269   |  |                                     |
| CC                    | COMPUTER READABLE FORM:   |  |                                     |
| CC                    | MEDIUM TYPE: 3M Double Density 5 1/4" diskette                          |  |                                     |
| CC                    | COMPUTER: Wang PC   |  |                                     |
| CC                    | OPERATING SYSTEM: MS DOS Version 3.20                                   |  |                                     |
| CC                    | SOFTWARE: Microsoft   |  |                                     |
| CC                    | CURRENT APPLICATION DATA:   |  |                                     |
| CC                    | APPLICATION NUMBER: US/07/865,662F                                      |  |                                     |
| CC                    | FILING DATE: 07 April, 1992   |  |                                     |
| CC                    | CLASSIFICATION: 435   |  |                                     |
| CC                    | PRIOR APPLICATION DATA:   |  |                                     |
| CC                    | APPLICATION NUMBER: 07/688,326  |  |                                     |
| CC                    | FILING DATE: 22 April 1991  |  |                                     |
| CC                    | APPLICATION NUMBER: 07/588,922  |  |                                     |
| CC                    | FILING DATE: 27 September 1990  |  |                                     |
| CC                    | APPLICATION NUMBER: 07/210,405  |  |                                     |
| CC                    | FILING DATE: 23 June 1988   |  |                                     |
| CC                    | APPLICATION NUMBER: US 07/130,529                                       |  |                                     |
| CC                    | FILING DATE: 9 December 1987  |  |                                     |
| CC                    | APPLICATION NUMBER: US 07/068,176                                       |  |                                     |
| CC                    | FILING DATE: 30 June 1987   |  |                                     |
| CC                    | ATTORNEY/AGENT INFORMATION:   |  |                                     |
| CC                    | NAME: Irons, Edward S.  |  |                                     |
| CC                    | REGISTRATION NUMBER: 16,541   |  |                                     |
| CC                    | REFERENCE/DOCKET NUMBER: No. 5451670e                                   |  |                                     |
| CC                    | TELECOMMUNICATION INFORMATION:  |  |                                     |
| CC                    | TELEPHONE: (202) 783-6040   |  |                                     |
| CC                    | TELEFAX: (202) 783-6031   |  |                                     |
| CC                    | TELEX: No. 5451670e   |  |                                     |
| CC                    | INFORMATION FOR SEQ ID NO: 13:  |  |                                     |
| CC                    | SEQUENCE CHARACTERISTICS:   |  |                                     |
| CC                    | LENGTH: 105   |  |                                     |
| CC                    | STRANDEDNESS: Double  |  |                                     |
| CC                    | TYPE: Nucleic Acid  |  |                                     |

```
CC      TOPOLOGY: Linear
CC      MOLECULE TYPE: DNA
CC      ORIGINAL SOURCE: Synthetically Prepared
CC      IMMEDIATE SOURCE: Synthetically Prepared
SQ      SEQUENCE 105 BP; 15 A; 0 C; 8 G; 1 T; 81 OTHER.

Query Match          1.8%; Score 32; DB 1; Length 105;
Best Local Similarity 16.5%; Pred.No. 3.80e-07;
Matches 15; Conservative 28; Mismatches 48; Indels 0; Gaps 0;

Db       9 AKNSNNKSNNAVKGNKAKSGKNKSANAVNNGWAKNSNNKSNMGTFRONKINRAA 68
Qy       976 agagattgctaacacatgtcgtaacaccacaggtaaaccacatttgcagtgcagaa 1035
         |: : :: |: : : |: : |: | : | : | :| : : : ||
Db       69 KNNNSGVADNKNNASNNYDNGSGVADNKNA 99
Qy       1036 ttttggaataatgacctgtcgtaacccagaatgaaccacatttcgaatcga 1066
         : :||: l : : : :|: : ||

RESULT 12
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC      AC xxxxxx
DT
DE Sequence 14, Application US/08232463
CC Sequence 14, Application US/08232463
CC Patent No. 5670367
CC GENERAL INFORMATION:
CC APPLICANT: DORNER, F.
CC APPLICANT: SCHEIFLINGER, E.
CC APPLICANT: FALKNER, F. G.
CC TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232,463
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/935,313
CC FILING DATE:
CC APPLICATION NUMBER: EP 91 114 300.6
CC FILING DATE: 26-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEX: (703)683-4109
CC TELETYPE: 899149
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
```

***This Page Blank (uspto)***

CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC IMMEDIATE SOURCE:  
CC CLONE: ptZqpt-fls  
SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match 1.7%; Score 30; DB 1; Length 7218;  
Best Local Similarity 1.3%; Pred. No. 7.26e-06;  
Matches 5; Conservative 197; Mismatches 172; Indels 0; Gaps 0;

[illegible]

|        |  |  |
|--------|--|--|
| RESULT | 13   |  |
| ID     | US-08-238-163-5  | STANDARD; DNA; UNC; 215 BP.            |
| AC     | xxxxxx   |  |
| DT     |  |  |
| DE     | Sequence 5,  | Application US/08238163                |
| CC     | Sequence 5,  | Application US/08238163                |
| CC     | Patent No.   | 5569830                                |
| CC     | GENERAL INFORMATION:   |  |
| CC     | APPLICANT:   | BENNETT, Alan                          |
| CC     | APPLICANT:   | LABAVITCH, John M.                     |
| CC     | APPLICANT:   | POWELL, Ann                            |
| CC     | APPLICANT:   | STOTZ, Henrik                          |
| CC     | TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL                                 |  |
| CC     | TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE |  |
| CC     | NUMBER OF SEQUENCES: 24  |  |
| CC     | CORRESPONDENCE ADDRESS:  |  |
| CC     | ADDRESSEE:   | Townsend and Townsend Khourie and Crew |
| CC     | STREET:  | Steuart Street Tower, One Market Plaza |
| CC     | CITY:  | San Francisco                          |
| CC     | STATE:   | California                             |
| CC     | COUNTRY:   | US                                     |
| CC     | ZIP:   | 94105-1493                             |
| CC     | COMPUTER READABLE FORM:  |  |
| CC     | MEDIUM TYPE: Floppy disk   |  |

```

CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc.feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of pcgip from bean."
CC SEQUENCE 215 BP: 15 A; 8 G; 25 C; 26 T; 141 OTHER.

```

[illegible]

RESULT 14  
US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.  
AC xxxxxx  
DT  
DE Sequence 5, Application US/08238163  
CC Sequence 5, Application US/08238163  
CC Patent No. 5569830  
CC GENERAL INFORMATION:  
CC APPLICANT: BENNETT, Alan  
CC APPLICANT: LABAVITCH, John M.  
CC APPLICANT: POWELL, Ann  
CC APPLICANT: STOTZ, Henrik  
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL  
CC TITLE OF INVENTION: POLYGLACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE  
CC NUMBER OF SEQUENCES: 24  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend Kourie and Crew  
CC STREET: Steuart Street Tower, One Market Plaza  
CC CITY: San Francisco

**This Page Blank (use)**



Sep 16 10:24

US-08-952-741-1 mi

23

CC STATE: California  
 CC COUNTRY: US  
 CC ZIP: 94105-1493  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/238,163  
 CC FILING DATE: 03-MAY-1994  
 CC CLASSIFICATION: 800  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Bastian, Kevin L.  
 CC REGISTRATION NUMBER: 34,774  
 CC REFERENCE/DOCKET NUMBER: 2307E-540  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (415) 543-9600  
 CC TELEFAX: (415) 543-5043  
 CC INFORMATION FOR SEQ ID NO: 5:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 215 base pairs  
 CC TYPE: nucleic acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: unknown  
 CC MOLECULE TYPE: protein  
 CC FEATURE:  
 CC NAME/KEY: misc feature  
 CC LOCATION: 1..215  
 CC OTHER INFORMATION: /standard name= "Deduced amino acid  
 CC OTHER INFORMATION: sequence of PCIP from bean."  
 CC SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 1.5%; Score 26; DB 1; Length 215;  
 Best Local Similarity 13.7%; Pred. No. 2.12e-03;  
 Matches 23; Conservative 61; Mismatches 84; Indels 0; Gaps 0;  
 Db 34 TDCQRTMGVCDTDTTYRVNDSCHNKYSSANNYGNVGAAKTHYTHNVSGADSKT 93  
 Cp 1091 ttccaactgtttttaaatttttcgatttcgacgaaggtcatttttccaaaattct 1032  
 Db 94 VTDSTNASGTSNNGDGRSGADSYGSKTAMTSRRTGKTANNVDSRNMGDASVGS 153  
 Cp 1031 gcaactgcaaacattggtttacctgtggttacgcacatgtgttagccaattctcgt 972  
 Db 154 DRNITKHKNSADGKVKSKNGDRNNRYGTCTKSNVNNCGGNKRDV 201  
 Cp 971 tagctgtatttaataatgtttcacagcatcgatcttaaatccatctaga 924

RESULT 15  
 ID PCT-US95-10426-3 STANDARD; DNA; UNC; 34 BP.  
 AC xxxxxx  
 DT  
 DE Sequence 3, Application PC/TUS9510426  
 CC Sequence 3, Application PC/TUS9510426  
 CC GENERAL INFORMATION:  
 CC APPLICANT: GENENCOR INTERNATIONAL, INC.  
 CC TITLE OF INVENTION: An Improved Cleaning Composition  
 CC NUMBER OF SEQUENCES: 68  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Genencor International  
 CC STREET: 180 Kimball Way  
 CC CITY: South San Francisco

Sep 16 10:24

US-08-952-741-1 mi

24

CC STATE: CA  
 CC COUNTRY: USA  
 CC ZIP: 94080  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/US95/10426  
 CC FILING DATE:  
 CC CLASSIFICATION:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: STONE, Christopher L.  
 CC REGISTRATION NUMBER: 33,401  
 CC REFERENCE/DOCKET NUMBER: GC220-3  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (415) 742-7536  
 CC TELEFAX: (415) 742-7217  
 CC INFORMATION FOR SEQ ID NO: 3:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 34 base pairs  
 CC TYPE: nucleic acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: DNA (genomic)  
 CC SEQUENCE 34 BP; 9 A; 7 C; 8 G; 10 T; 0 OTHER.

Query Match 1.4%; Score 24; DB 2; Length 34;  
 Best Local Similarity 85.3%; Pred. No. 3.17e-02;  
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Db 1 TCATGCAGTACTTTGATGCTTACCTGCCCAATGA 34  
 Cp 263 tgatgcagatatttgatgcagcatttgcccaatga 296

Search completed: Wed Sep 16 10:27:47 1998  
 Job time : 98 secs.

**This Page Blank (usp)**

\*\*\*\*\*

WIREH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Sep 16 08:46:42 1998; MasPar time 2402.17 Seconds  
Tabular output not generated. 1358.207 Million cell updates/sec

Title: >US-08-952-741-1  
Description: (1-1776) from US08952741.seq

Perfect Score: 1776  
N.A. Sequence: 1 atataaattgaatgaaca.....ccaataataattggaagcctt 1776  
Comp: tatattaaactttacttgt.....ggttatattacacctcgaa

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 532261 seqs, 918536377 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb155

1:em\_ba 2:em\_htg 3:em\_hum1 4:em\_hum2 5:em\_in 6:em\_om  
7:em\_or 8:em\_ov 9:em\_pat 10:em\_pl 11:em\_ro

Database: genbank107  
12:gb\_ba 13:gb\_htg 14:gb\_in 15:gb\_om 16:gb\_ov 17:gb\_pat  
18:gb\_ph 19:gb\_pl 20:gb\_pr1 21:gb\_pr2 22:gb\_ro 23:gb\_st  
24:gb\_sts 25:gb\_sy 26:gb\_un 27:gb\_vi

Statistics: Mean 11.347; Variance 5.064; scale 2.241

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No.              |
|------------|-------|-------------|--------|----|-------------|------------------------|
| 1          | 785   | 44.2        | 2397   | 12 | BACAMYG6    | Bacillus sp. (alkaloph |
| 2          | 473   | 26.6        | 2447   | 12 | BSU22045    | Bacillus sp. alpha-amy |
| 3          | 432   | 24.3        | 1990   | 12 | AF032864    | Bacillus stearothermop |
| 4          | 432   | 24.3        | 3048   | 12 | BACAMYLB    | Bacillus stearothermop |
| 5          | 430   | 24.2        | 1680   | 17 | E01181      | DNA encoding highly th |
| 6          | 430   | 24.2        | 1719   | 17 | E01180      | DNA encoding highly th |
| 7          | 430   | 24.2        | 2393   | 12 | BSU75445    | Bacillus sp. MK 716 a  |
| 8          | 428   | 24.1        | 1814   | 12 | BSAMYSEN    | B.stearothermophilus D |
| 9          | 428   | 24.1        | 1891   | 12 | BACAMYABS   | B.stearothermophilus a |
| 10         | 428   | 24.1        | 2169   | 12 | BSAMYLAI    | Bacillus stearothermop |
| 11         | 424   | 23.9        | 1650   | 17 | E01157      | DNA sequence of B.stea |
| 12         | 424   | 23.9        | 2066   | 12 | BACAMYS     | B.stearothermophilus a |
| 13         | 376   | 21.2        | 1539   | 17 | A47677      | Sequence 1 from Patent |
| 14         | 376   | 21.2        | 5677   | 17 | A21895      | Nucleotide sequence of |
| 15         | 374   | 21.1        | 1449   | 17 | A23402      | B.licheniformis gene f |

|    |     |      |      |    |           |                         |
|----|-----|------|------|----|-----------|-------------------------|
| 16 | 374 | 21.1 | 1449 | 17 | A27772    | Alpha amylase coding s  |
| 17 | 374 | 21.1 | 1777 | 17 | A17930    | Alpha amylase gene.     |
| 18 | 374 | 21.1 | 1777 | 17 | I24553    | Sequence 33 from paten  |
| 19 | 374 | 21.1 | 1948 | 12 | BACAMYLG  | Bacillus licheniformis  |
| 20 | 374 | 21.1 | 1948 | 12 | BACAMYABL | B.licheniformis alpha-  |
| 21 | 374 | 21.1 | 1968 | 12 | BACAMYS   | B.licheniformis amylase |
| 22 | 374 | 21.1 | 2149 | 17 | A21893    | Nucleotide sequence of  |
| 23 | 374 | 21.1 | 2829 | 17 | I08486    | Sequence 13 from Patent |
| 24 | 374 | 21.1 | 4190 | 17 | I08488    | Sequence 19 from Patent |
| 25 | 355 | 20.0 | 1536 | 17 | E01158    | DNA sequence of B.lich  |
| 26 | 352 | 19.8 | 1972 | 17 | A20154    | alpha-amylase gene (an  |
| 27 | 352 | 19.8 | 2084 | 12 | BACAM     | Bacillus amyloliquefac  |
| 28 | 158 | 8.9  | 2182 | 17 | I08183    | Sequence 1 from Patent  |
| 29 | 156 | 8.8  | 1900 | 12 | AF055987  | Streptococcus mutans i  |
| 30 | 141 | 7.9  | 2050 | 12 | BCAMYE    | B.circulans amylase     |
| 31 | 128 | 7.2  | 2302 | 12 | SBU04956  | Streptococcus bovis am  |
| 32 | 114 | 6.4  | 2300 | 12 | AB000830  | Streptococcus bovis ge  |
| 33 | 98  | 5.5  | 7218 | 17 | I66494    | Sequence 14 from paten  |
| 34 | 84  | 4.7  | 454  | 12 | BACAMYL   | B.licheniformis alpha-  |
| 35 | 84  | 4.7  | 454  | 17 | I00838    | Sequence 10 from Patent |
| 36 | 83  | 4.7  | 576  | 17 | A00157    | Nucleotide sequence 3   |
| 37 | 83  | 4.7  | 576  | 17 | A00604    | B.amyloliquefaciens al  |
| 38 | 83  | 4.7  | 576  | 17 | A00155    | Nucleotide sequence 1   |
| 39 | 83  | 4.7  | 576  | 17 | A00158    | Nucleotide sequence 4   |
| 40 | 83  | 4.7  | 576  | 17 | A00605    | B.amyloliquefaciens al  |
| 41 | 83  | 4.7  | 576  | 17 | A00156    | Nucleotide sequence 2   |
| 42 | 65  | 3.7  | 465  | 17 | E00040    | DNA coding of alpha-am  |
| 43 | 65  | 3.7  | 537  | 12 | BAMYL     | Bacillus amyloliquefac  |
| 44 | 63  | 3.5  | 465  | 17 | A00607    | B.amyloliquefaciens al  |
| 45 | 63  | 3.5  | 465  | 17 | A00608    | B.amyloliquefaciens al  |

## ALIGNMENTS

RESULT 1  
LOCUS BACAMYG6 2397 bp DNA BCT 15-MAR-1989  
DEFINITION Bacillus sp. (alkalophilic) G6-amylase gene, complete cds.  
ACCESSION M18862  
NID g142496  
KEYWORDS amylase; maltohexaose-producing amylase.  
SOURCE Bacillus sp. (alkalophilic strain #707) DNA, clone pTUE306.  
ORGANISM Eubacteria; Firmicutes; Low G+C gram-positive bacteria;  
Bacillaceae; Bacillus.  
REFERENCE 1 (bases 1 to 2397)  
AUTHORS Tsukamoto,A., Kimura,K., Ishii,Y., Takano,T. and Yamane,K.  
TITLE Nucleotide sequence of the maltohexaose-producing amylase gene from  
an alkalophilic Bacillus sp. #707 and structural similarity to  
liquefying type alpha-amylases  
JOURNAL Biochem. Biophys. Res. Commun. 151, 25-31 (1988)  
MEDLINE 88162814  
COMMENT Draft entry and computer-readable sequence for [1] kindly provided  
by K.Yamane, 02-JUN-1988.

FEATURES  
source  
1. .2397  
Location/Qualifiers  
/organism="Bacillus sp."  
/db\_xref="taxon:1409"  
729..827  
/note="G6-amylase signal peptide"  
729..2285  
/note="G6-amylase precursor"  
/codon\_start=1  
/transl\_table=11  
/db\_xref="PID:g142497"  
/translation="MKMRTGKGLSLLAFLLVITIPFTLVDEAHNNTGTFMMQ  
YFWMYLPNDGNHNWNRSLNKGITAVIPPAWKAGSQNDVGAYDLYLDGEF  
NQGVTRTYGTRSQAAVTSKNGIQVGVVMHKGADATMVAVEVNPNGR  
NQGVTEYTIETWTRDFPCRGTHSFQKRWYHFDGVDQSRLLNNRIYKFGHKG  
ANDWEVDENGNDYLMYADIDMDHPVWELRWGVTYNTLIDGFRIADVAHKY  
SFRDMNVRNSATGNMFAVAEKNLDGAIENLYQKNNHVSFVDFPLHNLNLS  
KSGNVDNRNIFNGTVVVRHSHAVTFVDNHDHSDQEALESFVEWFKPLAYLTLR  
EQGYPVFYGDYGIPTHPVPMRKIDPILQARKYAKQNDLYLHNTIGTREG  
NTAHPNSGLATIMSDGAGSKMFWGVRKAGQVMSDITGRTGTVTINADGWNFSVN

```
GGSVSIWVVK"
mat_peptide      828..2282
/notes="G6-amyase"
BASE COUNT      788 a 346 c 549 g 714 t
ORIGIN          1 bp upstream of BamHI site.

Query Match
Best Local Similarity 76.7%; Pred. No. 0.00e+00;
Matches 1128; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

Db 823 AACGACATCATACGGTACCAAGCGGACGATGATGCAATACCTTTGAATGGTATCTACCTA 882
QY 233 aagcccatcataatggagcgaatgggaccatgagtcgatttttgaatggcattgcaaa 292

Db 883 ATGACGGAATCATTTGGAATTCGATTAACCTCTGTAGCGAGTAACTTTAAAGCAAGGGA 942
QY 293 atgacgggaaccactggaaacaggttaacagatgacgcagctaaacttaaaagataaagga 352

Db 943 TTACAGGGGTGGATTCCTCCAGCATGGAAGGCGCTTCTCAAAATGACGTAGGATACG 1002
QY 353 ttaccgctgttgattccctccctgcatgggaagggaacttcgcaaaatgatgttgggtatg 412

Db 1003 GAGCCTATGACCTGTATGATCTGGGAGAATTAATCAAAAAGGTACCGTCGTACAAAT 1062
QY 413 gtcccatgattgtgcagatctgttgagtttaacccaaagggaacccgtccgtacaaaat 472

Db 1063 ATGGAACACGTAGTCTAGTTACAAAGCTCGGTAACTCTCTTAAATAATATGGAATTCAG 1122
QY 473 atggcaaggagtcagttgcaagggtgcgtgacatcttggaaaaataacgggattcaag 532

Db 1123 TATATGGTGACGTGTTTATGAATACAAAGGTGGCGGACGCTACTGAAATGTAAGGG 1182
QY 533 ttatggggatgctgtagatgaatacaaaaggtggagcagcgggacagagatggtaaatg 592

Db 1183 CGCTTGAAGTGAATCCCAATTAACCGTAACCAAGNAGTGTGATGATATACCAATTCAG 1242
QY 593 cgggtgaagtgaaccgaagcaacgaacaaagaaatcatcaggtggaatacaccattgaag 652

Db 1243 CTGGACTAGATTTGATTTCCAGGGCGAGGAAATACTCTAGCTTTAAATGGAGAT 1302
QY 653 catggcgaattgtatttcccttggaaagaaataacccattccaactttaaattggcgct 712

Db 1303 GGTATCATTTTGTGTTGGATGGGATCAGTCACGTAGCAGTAAATCGCATCTATA 1362
QY 713 ggtatcatttggatggagacagattgggatacagtcagtcagtcagtcagcaacaaatata 772

Db 1363 AATTTAGAGTCATGCAAGCTTGGGATGGGAAGTGTGATACGGAATATGTAATTAATG 1422
QY 773 aattcagaggtaccggaaaggcagtgaggactgggaagtagatatagagaacggcaactatg 832

Db 1423 ATTATTAAATGTACGCTGATATTGATGATGATCACCAGAGTAGTAATGAATTAAGAA 1482
QY 833 attaccttatgtatgcagacattgatggatcatccagaagtaatacaatgaacttgaag 892

Db 1483 ATTGGGGTGTGGTACAAACACATTAGACTCGATGATGATTAATAGATCGCGTTA 1542
QY 893 attggggagttgtgtatacaaaacacctaatactagatggatttagaatcgatgctgtga 952

Db 1543 AACATATAAGTATAGCTTTACGCGGATGGGATTAATCATCAGTGTAGAGTGCACAGGTA 1602
QY 953 aacataataatacagctatacagagatgtggctaacacacatgctgcgttaacacacaggt 1012

Db 1603 AAAATATGTTGGGTTGTGCTAGCTTTTGGGAAGATGATTTAGGTGCAATTTGAAAACATATC 1662
QY 1013 aaccaatgttgcagttgcagaattttggaaaaatgaacctgtgcaatgcaatgaactatt 1072

Db 1663 TGCAGAAAACAACTGGAACCATTCAGTCTTGTGATGTCGCGTGTACATTAATATCTTTATA 1722
QY 1073 taaataaaacaagtgtgaatcactccgtgttcgattgttcctcttcattataattgtaca 1132

Db 1723 ATGCATCAAAAAGCGGAGGAACATGATGATGGAACATATTTAATGGACCGTGTCTC 1782
QY 1133 atgcatactaagtggtgggctatttttgatagagaataattttcaaatggttctgtgtac 1192
```

---

```
1783 AACGACATCAAGTCATGCTGTAAACATTTGTTGATAATCATGATTCGACGCTGAAGAAG 1842
1193 aaaaacacctatacatcagtcacattgttgataccatgactctccgagcggaggaag 1352
1843 CATTAGAAATCTTTTGTGAAGAAATGGTTTAAACCATTTAGCGTATGCGCTTACATTAACGC 1902
1253 catggaaatcctttgttcaatcgtgtttcaaacccactgscatgcatgattctgacaa 1312
1903 GTGAACAGGATACCCCTCTCTGTATTATTACGGAGATTAATTATGGGATTCACACATGGAG 1962
1313 gggagcaaggttacccctccgtattttacggtgattactacggtataccaactcatggtg 1372
1963 TGCAGCAATGAGATCAAAAATTCGATTCGATTTTGAAGCAGCTCAAAAGTATGTCATACG 2022
1373 ttccctcgatgaatactcaaatgtatccactctcgcagggcacgtcaaacgtatgacctag 1432
2023 GAAACAAATGATTACTTAGACCATCATATAATCATTTGGTTGGAGCGCTGAAGGGAATA 2082
1433 gaaccaacatgatttttgcattcatgatattatcggtcgacgagagaaggagaca 1492
2083 CAGCACACCCCAATTCAGGCTAGCTACATCATGCTGTGATGAGCGGTGGAAGTAAAT 2142
1493 gctcccaaccacaaattcaggacttgcaactattatgctcgatggccaggggttaataat 1552
2143 GGATGTTTGGGCGTAAATAAGGCTGCTCAAGTATGGAGTATGAGTATTCACAGGAACCGTA 2202
1553 ggaatgtagtcgggaaacataaagcgtggcgaagtatgagagatatccaccggaatacaggt 1612
2203 CAGTACGGTTTACAAATCAATCAAGTCAGCGGTGGGCAATTTCTGTGAATGAGGCTCAG 2262
1613 ctggtacgcgtcaccattaatgcagatggttgggggaatttcactgtaaacgagggggcag 1672
2263 TTTCTATTGGGTCACAAATTAAGAGTGAA 2293
1673 ttccggttgggtgaagcaataataagaa 1703

RESULT 2
LOCUS      BSU22045      2447 bp      DNA      BCT      22-MAR-1995
DEFINITION Bacillus sp. alpha-amyase (Amy) gene, complete cds.
ACCESSION  U22045
NID         g722278
KEYWORDS   .
SOURCE      Bacillus sp. (TS-23).
ORGANISM    Bacillus TS-23
            Eubacteria; Firmicutes; Low G+C gram-positive bacteria;
            Bacillaceae; Bacillus.
REFERENCE   1 (bases 1 to 2447)
            Lin, L.-L., Chu, W.S. and Hsu, W.H.
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 2447)
            Lin, L.-L.
AUTHORS     Direct Submission
TITLE       Submitted (01-MAR-1995) Long-Liu Lin, Food Industry Research
JOURNAL     Institute, Culture Collection and Research Center, 331 Food Road,
            Hsinchu, Taiwan 300, Republic of China
FEATURES    Location/Qualifiers
            source          1..2447
                        /organism="Bacillus TS-23"
                        /db_xref="taxon:38441"
            source          1..2447
                        /organism="Bacillus sp."
                        /strain="TS-23"
                        /db_xref="taxon:1409"
            gene            241..2082
                        /gene="Amy"
            CDS             241..2082
                        /gene="Amy"
                        /EC_number="3.2.2.1"
                        /codon_start=1
                        /transl_table=11
                        /product="alpha-amyase"
```



```

/transl_table=11
/product="alpha amylase"
/db_xref="pid:g2642326"
/translation="MLTFHRIIRKGMWFLAFLTLASFCPTGOHAKAAPENGTMMO
YFENWLPDGLTWKTVANEANLSSGLTALWLPAYKGTSRSDVGYDYLDLGEF
NOKGTVRTKYCTKQYLOAQAHAGNOYADVDFDHKGADGTEWDAVEVNSDR
NQEISGTYQIOAWTKYDFDPRGNTYSEFKRWYHFDGVWDDESRLKSIYKFRGIGKA
WDEWYENGNYDILMYADLMDHPVEVTEKNGWKVYNTNIDGFRDLAVKHIFS
FFPDLWSYRSOTQKPLFTVGEYMSYDINKLHNYITKTNGTMSLFDAPLHKNFYTASK
SGGAFDMRTLMTIMKDOPTLAVTFVNDHDETPGOALQSWDFWFKPLAVAFILTRO
EGYPCVGYGIGYPOYNIPLSKIDPLLLIARROYAGTODHLDHSDIIGWTRGV
TEKPSGLAALITDPSGSKWYVKQHAGKVFVDLGNRSDTVIINDSGWGEFVNG
GSVSWVPRKTVTSIARPIITRWTFEVRWTEBRLVAMP"
BASE COUNT      529 a 496 c 488 g 477 t
ORIGIN

```

```

Query Match      24.3%; Score 432; DB 12; Length 1990;
Best Local Similarity 65.6%; Pred. No. 0.00e+00;
Matches 940; Conservative 0; Mismatches 490; Indels 3; Gaps 3;

Db 263 AACGGCACCATGATCCAGTATTTGAATGGTACTTTCGGGATGATGGACGTTATGGACC 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 253 aatggaccatgatgcagtatttgaatggcatttgcgaatgacgggaacacacggaac 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 AAAGTGGCCATGATGACCAACACTTATCCAGCTTGGCATCACCGCTCTTTGGGTGCGG 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 313 aggttacgagatgacgcagctaaacttaagagtaaggattaccgcgttttggattcc 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 CCCGCTTACAAAGGAACAGCGCAGCGATGAGGTACGGATGACGAGTATGATGAC 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 373 cctgcatgaaggagacttcgcaaatgatgttgggtatggcctatgatttgcacgat 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 CTCGGCGAATTCAATAAAGGACCTTCGCGACAAATATGGAACAAAGCTCAATAT 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 433 cttggtgagtttaaccaaaagggaacccgtccgtacaaataatgacacaggagtcagt 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 503 CTTCAAGCCATTCAAGCGCCGACCTGGAATGCAAGTGTACGCCGATGCTGTGTC 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 493 caaggtccgtgacatcttgaataataacgggattcaagtttatggggatgctgatg 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 563 GACCATAAAGCGCGCTGACGGCACGGAATGGGTGGACGCGTGAAGTCAATCCGTC 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 553 aatcataaaggtagcagacagagagatggttaaatgctggtggaagtgaaacgaagc 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 623 GACCGCAACCAAGAAATCTCGGACCTTATCAATCCAGCATGACGAAATTTGATTTT 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 613 aaccgaaaccaaataatcaggtgaaatacaccattgaagcagcagaatttgatttc 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 683 CCGGGCGGGCAACACTACTCTAGCTTTAAGTGGCGCTGGTACCATTGTCAGCGCGTT 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 673 cctggaagaggaaataccattcccaactttaaatggcgtggtatcattttgatgggaca 732
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 GATTGGGCAAGAACCCGAAAA-TTAAGC-CGCAAT-TACAAATTCGGGGCATCGGCAAA 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 733 gattgggatcagtcagtcagctcagacaaataataataatcagaggcaccggaag 792
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 800 CGGTGGATGGGAAGTAGACAGAAACGGAACACTATGACTACTTAAATGATCGCGAC 859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 793 gcatggagcgggaagttagatagatagaagaagcaactatgattcccttatgtatgcagac 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 860 CTTGATATGATATCCCGAAGTCGTACCGAGCTGTAACAACTGGGGGAATGTTATGTC 919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 853 attgatatgatatccagaagtaataatgaactgaagattggagtttggtatataca 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 920 AACACACGAACATTTGATGGGTTCGGCTTGATCGCTCAAGCATATTAAGTTCAGTTT 979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 913 aatacacttaactagatgatttagaatcgatcgtgaaacataataatacagcctat 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 980 TTTCCCTGATTTGTTGCTGATGTCGTTCTCAGACTGGCAAGCGCTATTATACGTCGGG 1039
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 973 acgagagattggttaacacacatgtgcgtaacacacaggttaacacacattgttgcagttgca 1032
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1040 GAATATTGGAGCTATGACATCAACAAGTTGCACATTTACATTACGAAAAACAACGGGACG 1099
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy 1033 gaattttgaaaaaatgaccttgcgtcaatcgaaaactattttaataaaacagttggaat 1092
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 ATGCTCTTTGTCATCCCGGTTACACACAAATTTTATACCGCTTCCAAATACAGGGGC 1159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1093 cactcgtgttcgatttctcttcataatttgacaaatgcataatagtggtgagc 1152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1160 GCATTGTATGGCAGCTTAATGACCAATACTCTCATGAAAGTCAACCGACATTCGGCC 1219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1153 tatttgatatagaaaatttttaaatggtctcgtacaaaacacccctatacatgca 1212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1220 GTCACCTTCGTTGATTAATGACACCGAACCGGCGCAACGCTGCAGTTCATGGTGCAC 1279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1213 gtcacatttggataaccatgactctcagcagagagacattggaatccttgttcaa 1272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1280 CCATGGTTCAACACCGTTGGCTTACGCTTTATTTCTAACCGGAGGAGATACCGCTGC 1339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1273 tegtgttcaaacacctggcatatgattgattctgacaaggaggaaggttacccttcc 1332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1340 GTCCTTTATGGTGACTATTATGCAATTCACCAATATAACATTCCTTCGCTGAAACGAAA 1399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1333 gattttacggtgattactacggtatcaccaactcaatggttctcttcgatgaaatctaa 1392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1400 ATCGATCCGCTCCTCATCGCGCAGGGATTATGCTTACGGACGCAACCAATGATTATCTT 1459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1393 attgatccacttctgcaggcagctcaaacgctatgcctacggaacccaacatgattttt 1452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1460 GATCACTCGGACATCATCGGGTGGACAGGAGGAGGCTCACTGAAACACAGGATCCGG 1519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1453 gatcatcatgatattatcggtggcagagagagggacagctccaccacaaatcaggga 1512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1520 CTGGCGGCACTGATCACCGATGGCGGAGGAGGAGCAAAATGGATGTACGTTGGCAACAA 1579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1513 cttgcaactattatgctcgatggcgagggggtataataatggtatgctcgggaaacat 1572
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1580 CACGCTGGAATAATGTTCTATGACCTTACCGGCAACCGGAGTGACACCGTCAACATCAAC 1639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1573 aaagctggccaagtatgagagatatccaccggaataatggtctggtaccgtccaccataat 1632
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1640 AGTGATGATGGGGGAATTCAAAGTCAATGCGGCTTCGGTTCGGTTTGGGT 1692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1633 gcagatggttgggggaatttcactgtaaacggagggcagtttcggttgggt 1685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 4
LOCUS      BACAMVLB      3048 bp      DNA      BCT      01-SEP-1988
DEFINITION Bacillus stearothermophilus alpha-amylase gene.
ACCESSION  M11450
            g142504
NID
KEYWORDS   Bacillus stearothermophilus DNA.
SOURCE     Bacillus stearothermophilus
            Eubacteria; Firmicutes; Low G+C gram-positive bacteria;
            Bacillaceae; Bacillus.
REFERENCE  1 (bases 1 to 3048)
            Nakajima,R., Imanaka,T. and Alba,S.
AUTHORS    Nucleotide sequence of the Bacillus stearothermophilus
TITLE      alpha-amylase gene
JOURNAL    J. Bacteriol. 163, 401-406 (1985)
MEDLINE    85234394
FEATURES   Location/Qualifiers
            1..3048
             /organism="Bacillus stearothermophilus"
             /db_xref="taxon:1422"
             680..2329
             /partial
             /note="alpha-amylase (gtg start codon; EC 3.2.1.1)"
             /codon_start=1
             /transl_table=11
             /db_xref="pid:g142505"
             /translation="VLTFHRIIRKGMWFLAFLTLALFCTGTPAKAAAPNGTMMO
             YFENWLPDGLTWKTVANEANLSSGLTALWLPAYKGTSRSDVGYDYLDLGEF
             NOKGAVRTKYCTKQYLOAQAHAGNOYADVDFDHKGADGTEWDAVEVNSDR"

```



```
/organism="Bacillus stearothermophilus"
/db_xref="taxon:1422"
BASE COUNT      450 a      406 c      425 g      399 t
ORIGIN

Query Match      24.2%; Score 430; DB 17; Length 1680;
Best Local Similarity 65.5%; Pred. No. 0.00e+00;
Matches 939; Conservative 0; Mismatches 491; Indels 3; Gaps 3;

Db 145 AACGGCCATGATGACGATTTTGAATGGTACTTTGCGGATGATGGCAGTTATGGACG 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 253 aatgggaccatgatcgatatttgaatggcatttgcacaaatgaagggaacacattggaac 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 AAATGGCCATGAAGCAACAACTATTCCAGCCTTGGCATCCGCTCTTTTGGGTGGCG 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 313 aggttacgagatgcgcgactaaacttaagagtaaggaagattaccgcgttttggattcct 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 265 CCCGCTTATAAGGAACAAAGCGCAGCGCTAGGCTAGCGAGTATACGACTTGTATGAC 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 373 cctgcattgaagggaacttcgcaaatgatgttgggtatggtgcctatgattgttacgat 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 CTCGGTGAATCAATCAAAAGGGCGCTCCGCACAAATACGGAACAAAGCTCAATAT 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 433 ctgttgagtttaaccacaaagggaaccgtccgtacaaatatggcacaaggagtcagttg 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 CTTCAAGCCATTCAAGCGCCGCCACCGCTGGAATGCAATGCAAGTGTACGCCGATGCTG 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 493 caaggtgcgtgacatctttgaaaaataacgggattcaagtttatggggtcgtgatg 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 GACCATAAAGCGCGCGCAGCGACGAATGGGTGGAGCGCTGCAAGTCAATCCGTC 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 553 aatcataaaggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 GACCGCAACCAAGAAATCTCGGCGACCTATCAAAATCCAAAGCATGGACGAAATTTGATTT 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 613 aacgcaaaccaagaataatcaggigaaacacattgagcagcagcagcagcagcagcagcagc 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 CCCGGCGGGCAACACCTACTCTCAGCTTTAAGTGGCGCTGGTACCATTTTATGGCGCT 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 673 cctggaagaggaataaccattcccaactttaaattgagcgtggtatcatttggatggaca 732
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 GATTGGGACGAAGACCGGAAAA-TTGAGC-CGCATT-TACAATTCGCGGGATCGGCAAA 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 733 gattgggatacagtcacgtcagctcagcagcagcagcagcagcagcagcagcagcagcagc 792
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 CGTGGGATTGGGAAGTAGACACCGGAAACGAAACTATGACTACTTAATGATGCGCGAC 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 793 gcatgggactgggaagtagatagatagagaacggaacgaactatgattacattatgtagcagac 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 742 CTTGATATGATATCCCGAAGTCTGACTGAGCTGAAAAGCTGGGGGAAATGGTATGTC 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 853 attgatggatcatccagagaagtaacaaatgaactagaattggggagtttgggtatata 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 802 AACACACGAACATATGATGGGTTCCGGCTTATGCGGTCAGCAAGCATATTAAGTTTCAGTTT 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 913 aatacacttaactagatgagatttagaatcagatcgtggaacataattaaatacagcact 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 862 TTTCCCTGATTGGTTGTCGAATGTCGTTCTCAGACTGGCAAGCGCTATTTACCGTTGGG 921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 973 acgagagattggctaacaacatgctgtaaacaccacaggtaaacaaattgttgcagttgca 1032
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 922 GAATATGGAGCTATGACATCAACAAAGTTGCAACATTTACATTTATGAACAAACGGAACG 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1033 gaattttggaaaaatgaccttgcgtgaatcgaaacactattttaaataaaacaaagttggaat 1092
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 982 ATGTCTTTGTTGATGCGCGGTTACACAAACAAATTTATACCGCTTCCAAATCAGGGGGC 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1093 cactccggttcgaagtcttcattataattgttacaatgcatcctaagtgtggtg 1152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1042 ACATTTGATATGCGCAGCTTAATGACCAATCTCTATGAAGATCAACCAACATGCGCC 1101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1153 tattttgatagagaaatattttaaattgggttctgctgtaaaaaaacacccctatacatgca 1212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



|                       |        |                     |        |                                   |
|-----------------------|--------|---------------------|--------|-----------------------------------|
| Query Match           | 24.2%  | Score 430;          | DB 17; | Length 1719;                      |
| Best Local Similarity | 65.5%; | Pred. No. 0.00e+00; |        |                                   |
| Matches               | 939;   | Conservative        | 0;     | Mismatches 491; Indels 3; Gaps 3; |

|    |      |  |      |
|----|------|--|------|
| Db | 184  | AACGGCACCATGATGCAGTATTTTGAATGTGACTTGCCGGATGATGCGACGTTATGGACC       | 243  |
| QY | 253  | aa599gaccatgacgcagtgattttgaaatggcatttgcgaatgacg99gaaccactggaac     | 312  |
| Db | 244  | AAAGTGGCCCAATGAAGCCAAACAATTCATCGACGCTTGGCATCCGCTCTTTGGCTGCC        | 303  |
| QY | 313  | aggttacgagatgacgcagcactaaagatgaaggaagattaccgctgtttgagattcct        | 372  |
| Db | 304  | CCCGCTTATAAGGAACAAGCCGACGCTGAGGTACGGAGTATACGACTTGTATGAC            | 363  |
| QY | 373  | ctcgtatggaagg99gacttcgcgaatgatgttgggtatgctgctatgatttgcagt          | 432  |
| Db | 364  | CTCGGTCAATTCATCAAAAAGGGCGCTCGGCACAAAATACGGAACAAAAGTCAATAT          | 423  |
| QY | 433  | ctt599tgagtttaaccaaaagg9gaaccgctccgtacaaaatat99cacaa9ggtcaggtg     | 492  |
| Db | 424  | CTTCAAGCGATTCAAGCGCGCCACGCGCTGGAAATGCAAGTGTACGCGGATGTCGTTTC        | 483  |
| QY | 493  | caa9gtccgtgacatctttgaaaaataaacgggattccaagtttatgg99gatgcgtgatg      | 552  |
| Db | 484  | GACCATAAAGCGCGCGACGCGACGCAAGTGGGTGACGCGCTGCAAGTCAATCCGTC           | 543  |
| QY | 553  | aatcataaggtggagcagcgggacagagatg9taaatcggttgg9aagtfgaac9gaagc       | 612  |
| Db | 544  | GACCGCAACCAAGAAATCTCGGCACCTATCAATCCAAGCATTGGAGCAAAATTGATTT         | 603  |
| QY | 613  | aac9gaaccaagaaatcatcaggtgaaatacaccattgaa9catg99gcgaatttgatttc      | 672  |
| Db | 604  | CCCGCGGGGCAACACTACTCCAGCTTTAAAGTGGCGCTGGTACCATTTTGATGGCGTT         | 663  |
| QY | 673  | cct9gaagagaaataaccattccaactttaaatg99cgtgggtatcatatttgatg99gaca     | 732  |
| Db | 664  | GATTGGGACGAAGCCGAAAAA - TTGAGC - CGCAATT - TACAAATTCGCGGCATCGGCAAA | 720  |
| QY | 733  | gat599gatcagtcacgtcagcttcagaaacaaataataaattcagaggt9aac9gaag        | 792  |
| Db | 721  | CGTGGGATTGGGAAGTAGACACGGAAACGGGAACATGACATCTTAATGTATCGCGAC          | 780  |
| QY | 793  | 9ca599gactg99gaagt9agatagaa9ga99caactatgattacccttatgtatgc9ac       | 852  |
| Db | 781  | CTTGATATGGATCATCCGGAAGTCGTGACATGAGCTGAAAGCTGGGGGAAATGGTATGTC       | 840  |
| QY | 853  | attgatattgataatccagaagtaatacaatgaactagaataattg999gagttt99gtataca   | 912  |
| Db | 841  | AACACAACGAACATTGATGGGTTCGCGCTTGATGCCCTCAAGCATATTAAGTTCAGTTTT       | 900  |
| QY | 913  | aatacacttaactagatggtattgaaatcgatcgtgtgaaacataataaatacagctat        | 972  |
| Db | 901  | TTTCCCTGATTTGGTTCGATGTCGGTCTTCAGACTGGCAAGCGGCTATTTTACCGTTGGG       | 960  |
| QY | 973  | ac9agagat1ggctaaacacatgtcgt9aaacacaggt9aaacaaatggttgcagttgca       | 1032 |
| Db | 961  | GAATATTGGAGCTATGACATCAACAAGTTGCACAAATTCATATGAAACAAACGGAACG         | 1020 |
| QY | 1033 | gaa4ttt999aaaaatgacctgtgcaatgc9aaacattttaataaacaagaattg9aat        | 1092 |
| Db | 1021 | ATGCTTTTGGTTGATGCCCGCTTACACAACAAATTTTATACCGCTTCCAAATCAGGGGGC       | 1080 |
| QY | 1093 | cactcoggttgcagtggttcctcttcattataatttgacaatgcatc9aatagttggtg9c      | 1152 |
| Db | 1081 | ACATTGTGATTCGCGAGCTTAAATGACCAATFACTCTCATGAAGATCAACCAACATTTGGCC     | 1140 |
| QY | 1153 | tattttgatatgagaaatatttttaaatggttctgtcgt9caaaaaacccctatacatgca      | 1212 |
| Db | 1141 | GTCACTTCGTTGATATCATGACACGAAACCCCGCCACGGCTGCAGTCATGGTGCAC           | 1200 |
| QY | 1213 | gtcacattgttgataaccatgactctcagcag99gaagcat1gg9aatcctt1ggttcaa       | 1272 |
| Db | 1201 | CCATGGTTCAACCGTTGGCTTACCGCTTTATTCTAATTCGGCAGGAAGATACCCGTGC         | 1260 |

[illegible]

|    |      |  |      |
|----|------|--|------|
| Db | 510  | CTTCAAGCCATTCAAGCCGCCACCGCGTGGAAATGCAAGTGTACGCCGATGCTGGTTC     | 559  |
| Qy | 493  | caaggtgcctgacatctttgaaaaataacgggttaoagtttatggggtatcgctgatg     | 552  |
| Db | 570  | GACCAATAAGGGGGCGCTCAGCGCAGCAATGGGTGGACGCGTGAAGTCATCCGTC        | 629  |
| Qy | 553  | aaTcaTaaagggTggagcagcgagagatggTaaatgcggTggagTgaaccgaagc        | 612  |
| Db | 630  | GACGCAACCAAGAAATCTCGGCGCACCTATCAAAATCCAAGCATGGAGCAAAATTTAT     | 689  |
| Qy | 613  | aaccgaaccaagaataTcaggtgaatacaccattgaagcatggacgaaattgatttc      | 672  |
| Db | 690  | CCGGGGGGGCAACACTTACTTCAGCTTTTAAGTGGCGCTGGTACATTTTGAAGCGGTT     | 749  |
| Qy | 673  | ctTggaagggaaatacccatTccaaCTTaaatgcgttggtatcatctttgatggaca      | 732  |
| Db | 750  | GATTGGGACGAAGCGGAAA--TTGAGC--CGCATT-TACAATTTCCGCGCATCGGAAA     | 806  |
| Qy | 733  | gaTgggatacgtTcagtcagctcagcttcagacaacaaataataatTcagaggTcacggaa  | 792  |
| Db | 807  | GCGTGGGATTGGGAAGTAGACACACGGAACAACTATGACTACTTAAATGATGCGGAC      | 866  |
| Qy | 793  | gcattggactgggaagtagatatagaaacgcgaactatgattaccattatgtatgcagac   | 852  |
| Db | 867  | CTTGATATGATATCCCGAATCGTGACCGAGCTGAAAACTGGGGAAATGGTATGTC        | 926  |
| Qy | 853  | atTgatggatcatccagaaTaatcaatgaacttagaaattggggagtttggtataca      | 912  |
| Db | 927  | AACACAACGAACATTGATGGGTTCGGCTTGATGCCGTCAAGCATATTAAGTTTCAGTTT    | 986  |
| Qy | 913  | aaTaaactTaatctagatggattagaatcgatcgtTgaacataTtaatacagcatat      | 972  |
| Db | 987  | TTTCCTGATTGGTGTGCTATGGCTGCTCAGACTGCACAGCGCTATTTCACGTCGGG       | 1046 |
| Qy | 973  | acgagagattgctaacaacatgtgctaaccaccacaggtaaacaaatgttgcagtTgca    | 1032 |
| Db | 1047 | GAATATGGAGCTATGACATCAACAAAGTTGCACAATTTACATTACGAATAACAGCGAACG   | 1106 |
| Qy | 1033 | gaattttggaaaaatgaaccttgctgcaatcgaaaaactattTaaataaaaaaagttggaat | 1092 |
| Db | 1107 | ATGCTCTTTGATGACCGCTTACACAACAATTTTATACGCTTCCAATACAGGGGGC        | 1166 |
| Qy | 1093 | caTccggtgtogattgtccctctcatataattTgtcaatgcataatTaatTggtTggc     | 1152 |
| Db | 1167 | GCATTGTATGGCAGCTTAATGACCAATCTCTCAAGATCAACCGACATTCGGCC          | 1226 |
| Qy | 1153 | TattttgatatagaaataattTaaatggtctgTcgTacaataaacacctatacatTgca    | 1212 |
| Db | 1227 | GTCACTTCGTTGATTAATCATGACACCGAACCGGCCAGCGCTGCAGTCTATGGGTGCAC    | 1286 |
| Qy | 1213 | gtcaatttgtataaccatgactctcagccagagagcatTggaatcctttgtTcaa        | 1272 |
| Db | 1287 | CCATGGTTCAACCGTTGGCTTACGCCCTTTATCTTAACTCGGCAGGAGATACCGGTGC     | 1346 |
| Qy | 1273 | tcgtggtTcaaacaccatggcatatgcattgtattctgacaaggaggaaggtTtaaccttc  | 1332 |
| Db | 1347 | GTCTTTTATGCTACTTATATGGCATTCACAATATAACATTTCTCGTCAAAAGCAAA       | 1406 |
| Qy | 1333 | gtattttacgggtatcactacaggtataccaaactcatgTgtctctctgatagaactcaaa  | 1392 |
| Db | 1407 | ATCGATCCGCTCTCATCGCGGCGAGGATTTATGTTTACGGAACGCAACATGATTATCTT    | 1466 |
| Qy | 1393 | atTgatccactctgcaggcagctcaaacgtatgcctacggaaccacaatgattatttt     | 1452 |
| Db | 1467 | GATCACTCCGACATATCGGGTGGACAGGGGAAGGGGCGACTGAAAAACAGAGATCCGGA    | 1526 |
| Qy | 1453 | gatcatatgatattcggctTggaagagagaggggagcagctccaccaccaaatTtcagga   | 1512 |
| Db | 1527 | CTGGCGCACTGATACCGATGGCGGGGAGGAACCAATGGATGTACTGTTGGCAACAA       | 1586 |
| Qy | 1513 | ctTgcaactattatgtTccgatgggcgggggTaaataaatTggaatgtatTgTcgggaacat | 1572 |
| Db | 1587 | CACGCTGGAATAAGTTCATGACCTTACCGGCAACCGGAGTGACACCGTCAACCATCAAC    | 1646 |

[illegible]

Db 409 AAGTGCCCAATGAAGCCACAACTATTACAGCCTTGCCATCACCCTCTTTGGTGCCG 468  
Qy 313 aggttacgagatgacgagctaaactaaagagtaaaagattaccgtgttggtatccct 372  
Db 469 CCGCTTTATAAGGAACAACCGCAGCGAGCTAGGGTACGGAGTATACGACTTGTATGAC 528  
Qy 373 cctgcatggaaggggacttcgcaaaatgagtgttggtgctgctgctgattgtaagat 432  
Db 529 CTCGGTGAATCAATCAAAAAGGGCGCTCCGCACAAAATACGGAACAAAGCTCAATAT 588  
Qy 433 ctgggtgagtttaacaaaggaagacgcgtccgtacaaaatggtcacaagagatcagttg 492  
Db 589 CTTCAAGCCATTCAAGCGCCCGCAGCGCTGGAATGCAAGTGTACGCCGATGTCGTTC 648  
Qy 493 caaggtgcgtgacatctttgaaaaataacacgggattcaagtttatggggtgctggtgatg 552  
Db 649 GACCATAAAGCGCGCGCGACGCGACGGAATGGTGTGACGCCGCTCGAAGTCAATCCGTC 708  
Qy 553 aatcataaaggttgagcagacgagacggtgagatggtcaatcggtggaagtgaaccgaagc 612  
Db 709 GACCGCAACCAAGAAATCTCGGCATCTATCAATCAAGCATGCGAGAAATTTGATTTT 768  
Qy 613 aaccgaaccaaagaatatcaggtgaatcacacccattgaagcatgagcaaaatttgatttc 672  
Db 769 CCGCGCGGGGCAACACCTACTCCAGCTTTAAGTGGCGCTGGTACCAATTTTGATGGCGTT 828  
Qy 673 cctggaagaggaataaccctatcccaacttcaactttaaattgagcgtggtatcatttggagaca 732  
Db 829 GATTGGGAGCAAGCGCAAAA-TTGAGC-OGCATT-TACAAATTCGCGCATCGCAAAA 885  
Qy 733 gattggatcagtcacgtcagcttcagacaaataataataataatcagaggtgagaaag 792  
Db 886 GCGTGGATGGGAAGTAGACACGGAACGGAACCTATGACTACTTAATGTATGCGGAC 945  
Qy 793 gcatggactgggaagtagatataagaaacggcaactatgattacaccttatgtatgcagac 852  
Db 946 CTGTGATGATGATCCCGAAGTCGTGACTGACCTAAAGCTGGGCGTAAATGGTATGTC 1005  
Qy 853 atgcatgagatcgcatcagaaagtaataatgaacttagaaattgggagttggtataca 912  
Db 1006 AACACAAGAACACTTGATGGTTCGCGCTGTGTCGCTCAAGCATATTAAGTTACGTTTT 1065  
Qy 913 aatacaactaatcagatgattagaatcagatcgtgtaaacattttaaataacacgctat 972  
Db 1066 TTTCCTGATGTTGTCGTATGTCGTCTCAGACTGGCAAGCGCTATTTACCGTTGGG 1125  
Qy 973 acgagagattggctaacaacatgctgtaaacacacaggtaaaccaaattgttcagttgca 1032  
Db 1126 GAATATTGAGCTATGACATCAACAGTTGCACAATTACATTATGAACAAACGGAACG 1185  
Qy 1033 gaatttggaaaatgacacttgcgtgcaatcgaaaactatttaataaaaacaagttggaat 1092  
Db 1186 ATGTCTTTTGTGATGCCCGCTTACACAAATAATTTATACCGCTTCCAAATCAGGGGC 1245  
Qy 1093 cactcgtgttcgattgctcttcaattataattgtacaatgcatactaatagtggtggc 1152  
Db 1246 ACATTGATATCGGACGTTAAGACCAATATCTCATGAAAGATCAACCAACATTGGCC 1305  
Qy 1153 tattttgatgagaaatattttaaattggttctgtctgtaaaaaaacacctatacatgca 1212  
Db 1306 GTCACCTGCTTGATTAATCATCACACCGAACCGCCACAGCGCTGCAGTCATGGGTGCAC 1365  
Qy 1213 gtcaatttgttgaataacatgactctcagcagagagaagcattggaatcccttgttcaa 1272  
Db 1366 COATGGTTCAACCGTTACGCTTTACGCTTTTATCTTAACCTCGGAGGAAGTACCGGTGC 1425  
Qy 1273 tegtgtgtcaaacacctggcatgattgattcagaaagggagcaaggttacccttcc 1332  
Db 1426 GTCCTTTATGGTACTATATGCGAATTCACAATATACATTTCCTTCGTGAAAAACAAA 1485  
Qy 1333 gtatttccggtgattactacggtgataccaactcattggttcttccttcgataaataaa 1392

Db 1486 ATCGATCCGCTCCATCCGCGCAGGATTATGCTTACGGAACCAACATGATTATCTT 1545  
Qy 1393 attaatccattctgcagcagctaaacgctatgctcaggaaccccaacatgattatttt 1452  
Db 1546 GATCATTCCGACATCATCGGTGGACAAAGGAAGGGTCACTGAAAAACCCAGGATCCGGA 1605  
Qy 1453 gatcatgatattatcggtgctgacgagagagagggagcagctccacccaattcagga 1512  
Db 1606 CTGCGCCGATTGATCACCAGTGGCGGGAGAGAACCAATAGATGATGCTTGGCAACAA 1665  
Qy 1513 cttgcaactattatgcogattgcccaggggtaataaatgattgattgctgggaaacat 1572  
Db 1666 CACCCCGGAAAAGTCTTCTATGACTTACCGGCAACCGAGTGACACGTCACCATCAAC 1725  
Qy 1573 aaagctggccaagatgagagatatccacggaaaataggtctggtaccgtccaccattaat 1632  
Db 1726 AGTATGATGGGGGAATTCAAAGTCAATGCGGCTTCGGTTTCGGTTTGGGT 1778  
Qy 1633 gcagatggttgggggaatttcactgtaaacggagggcgagttcgttgggttgggt 1685

## RESULT 10

LOCUS BSAMYL1 2169 bp DNA BCT 21-MAR-1995  
DEFINITION Bacillus stearothermophilus gene for alpha-amylase.  
ACCESSION X02769  
NID g39799  
KEYWORDS amylase; amylase-alpha; inverted repeat; signal peptide.  
SOURCE Bacillus stearothermophilus.  
ORGANISM Bacillus stearothermophilus  
Eubacteria; Firmicutes; Low G+C gram-positive bacteria;  
Bacillaceae; Bacillus.  
REFERENCE 1 (bases 1 to 2169)  
AUTHORS Ihara,H., Sasaki,T., Tsuboi,A., Yamagata,H., Tsukagoshi,N. and Uda,K.S.  
TITLE Complete nucleotide sequence of a thermophilic alpha-amylase gene: homology between prokaryotic and eukaryotic alpha-amylases at the active sites  
JOURNAL J. Biochem. 98 (1), 95-103 (1985)  
MEDLINE 86008166  
REFERENCE 2 (bases 1 to 2169)  
AUTHORS Tsukagoshi,N.  
TITLE Direct Submission  
JOURNAL Submitted (03-SEP-1985) to the EMBL/GenBank/DBJ databases  
COMMENT Data kindly reviewed (03-SEP-1985) by Tsukagoshi N.  
FEATURES  
Location/Qualifiers  
source 1..2169  
/organism="Bacillus stearothermophilus"  
/db\_xref="taxon:1422"  
repeat\_unit 66..74  
/note="inverted repeat a"  
repeat\_unit 79..87  
/note="inverted repeat a"  
repeat\_unit 142..150  
/note="inverted repeat b"  
repeat\_unit 161..169  
/note="inverted repeat b"  
promoter 209..214  
/note="put. -35 region"  
promoter 233..238  
/note="put. -10 region"  
RBS 253..256  
/note="pot. Shine-Dalgarno region"  
RBS 306..310  
/note="pot. Shine-Dalgarno region"  
RBS 338..343  
/note="pot. Shine-Dalgarno region"  
RBS 345..348  
/note="pot. Shine-Dalgarno region"  
CDS 354..2003  
/note="put. precursor"  
/codon\_start=1  
/transl\_table=11  
/db\_xref="FID:g580825"



[illegible]





KLINGTVSKHPLKSVTFVNDHDTQPOQSLSTVQTFWFKPLAYAFILTRSGYPOVFE  
GDMTGTGDSQREIPALKHKIEPIKARKOYAYCAOHDFDHDHIDVIGWTRGDSVAN  
SGLAALITDGPGGAKRMVYGRNAGETWHDITGNRSPVINSSEGWEFHVNGSVSI  
YVQR"

mat\_peptide

BASE COUNT 427 a 329 c 403 g 380 t

ORIGIN /product="ALPHA-AMYLASE"

Query Match 21.2%; Score 376; DB 17; Length 1539;  
Best Local Similarity 66.2%; Pred. No. 0.00e+00;  
Matches 952; Conservative 0; Mismatches 468; Indels 18; Gaps 9;

|    |      |  |      |
|----|------|--|------|
| Db | 97   | AATGGGACGCTGATGTCAGTATTITGAATGGTACATGCCAATGACGGCCCAACATTTGGAAG | 156  |
| Qy | 253  | aatgggacatgatgcagtgatttggaatggcatttgccaatgacgggaaccactggaac    | 312  |
| Db | 157  | CGTTTGCAAAACGACTCGGCATATTGGCTGAACACGGTATTACTGCCGCTCGATTCOC     | 216  |
| Qy | 313  | aggttacgagatgacgcagctaaactaaagagtaaaagggattaccctgttggattcct    | 372  |
| Db | 217  | CCGCGATATAAGGAAGTACTAGTCAAGCGGATGTGGCTACGGTCTTACGACCTTTATGAT   | 276  |
| Qy | 373  | ccgcgatgaaaggggacttcgcaaatgatgttggatgttggctatattgtacgat        | 432  |
| Db | 277  | TTAGGGGAGTTTCATCAAAAGGACGGTTCGGACAAAGTACGGCACAAAAGGAGAGCTG     | 336  |
| Qy | 433  | cttggtagtttaacaaaagggaacgctccgtacaaaataatggcaacaggagtcagttg    | 492  |
| Db | 337  | CAATCTGCCATCAAAAGTCTTCATTCGCCGACATTAACTGTTTACGGGATGTGTCATC     | 396  |
| Qy | 493  | caaggtagctgacatcttggaaaaataacgggattcaagtttatgggagtagctgag      | 552  |
| Db | 397  | AAACCAAAAGGCGGCTGATCGACCGCAAGATGTAACCGCGGTGGAAGTCGATCCCGCT     | 456  |
| Qy | 553  | aatcataaaggtaggcagacgagggagagatggttaaatgaggtggaagtgaacgaagc    | 612  |
| Db | 457  | GACCCCAACCGGTAAATTCAGGAGAACACCTTAATTAAGCCTGACACATTTTCATTTT     | 516  |
| Qy | 613  | aaccgaacccaagaataatcagtgtaataacacattgaagcatgagcaaatattgattc    | 672  |
| Db | 517  | CGGCGCGCGGACGACATACAGCGATTTAAATGGCATGTGTACCATTTTGACGGAAAC      | 576  |
| Qy | 673  | cctggaagagaataaccattccaaacttaaatggcgctggtatcattttgagggaca      | 732  |
| Db | 577  | GATGGGACGAGTCCGAAAGCT---GAACCGCATCTATAAGTTTC-A---A---GGAAG     | 627  |
| Qy | 733  | gattgggacagtcagtcagcttcagacaacaaataataaattcagaggtaccggaaag     | 792  |
| Db | 628  | GCTTGGGATGCGGAGTTTCCATGAACACGCAACTATGATTATTTGATGTCGCGAC        | 687  |
| Qy | 793  | gcattgggactgggaagtagatagatagagacggcaactatgattaccattatgtgcag    | 852  |
| Db | 588  | ATCGATTATGACCATCTCTGATGTCGACGAGAAATTAAGATGCGGCATTTGGTATGCC     | 747  |
| Qy | 853  | attgatattgatcatccagaagtaataatgaacttagaaatgggagtttggatata       | 912  |
| Db | 748  | AATGAACGCAATGGAGGTTTCGCTGCTGCTCAACACATTAATAATTTCTTTT           | 807  |
| Qy | 913  | aatacattatctagatgatttagaatcgatgctgtgaaacatatataaatacaactat     | 972  |
| Db | 808  | TTGCGGATTGGGTTTATCATGTTCAGGGAACAAACCGGAGGAAGATTTTACGGTAGCT     | 867  |
| Qy | 973  | acgagagattggctaacaacatgctgtaaccacacagggtaaaccaatgttgcagttga    | 1032 |
| Db | 868  | GAATATGCGAGAAATGACTTTGGCGCCCTGGAAACATTTTGAACAAAACAAATTTAAT     | 927  |
| Qy | 1033 | gaatttggaaaatgacctgtgcaatcgaaaactatttaataaaaacaggttggaaat      | 1092 |
| Db | 928  | CATTCAGTGTTCACGTGCGCTTCATTATTCATTCCTGCTGCAACACGAGGAGCG         | 987  |
| Qy | 1093 | cactcgtgttcgattctctcttcaattataattgtacaatgcattcaatagtggtgac     | 1152 |

|    |      |  |      |
|----|------|--|------|
| Db | 988  | GGCTATGATATGAGAAATTCGTGAACCGTAGCTACGTCGTTCCRAAGCATCCGTTGAAATCG | 1047 |
| Qy | 1153 | tatttgatagataaataatttcaaatggtctgtcgcaacaacacccatatacatgca      | 1212 |
| Db | 1048 | GTTACATTTGTCATACCATGATACACAGCCGGGCAATCGCTTGTAGTACGATGTGCA      | 1107 |
| Qy | 1213 | gtcacatttgtataacatgactctcagccagagagcattggaactcttgttcaa         | 1272 |
| Db | 1108 | ACATGGTTTAAAGCGCTGCTTACGCTTATTCCTACAGGGAATCTGGATACCTCTAG       | 1167 |
| Qy | 1273 | tcgtggttcaaacactcgcataatgattctcagaaaggagagcaaggttacccttcc      | 1332 |
| Db | 1168 | GTTTCTACGGGATATGATACGGGACGAAAGGAGATCCACGCGCAAAATTCCTGCTTG      | 1227 |
| Qy | 1333 | gtatttaccgtgattactacggtat---acca-actcat-g-gtg---ttccttcga      | 1383 |
| Db | 1228 | AAACCAAAATTTGAACCGATCTTAAAGCGGAGAAACAGATATGCGTACGAGCACAGCAT    | 1287 |
| Qy | 1384 | aaatcaaatatgataccactctcagcagcagctcaaacgtatgctacggaacccaacat    | 1443 |
| Db | 1288 | GATTATTCGACCCATGACATTCGCTGGACAAAGGAGCGACAGCTCGGTTGCA           | 1347 |
| Qy | 1444 | gattatttgcatacatgattatcgcgtgagcagagagaggggagcagctcccccaca      | 1503 |
| Db | 1348 | AATTCAGGTTTGGCGCATTAATAACAGACGACCCGCGTGGGCAAGCGAATGTATGTC      | 1407 |
| Qy | 1504 | aattcagagactgcaactattatgctcagatggcagggggtgaaataaatgagtgatg     | 1563 |
| Db | 1408 | GCCGCGCAAAACCGCGGTGAGACATGACATGACATTACCGGAAACCGTTCGAGCCGGTT    | 1467 |
| Qy | 1564 | gggaacataaagctggccaagtagtgagagatatccacggaaatagtgctggtacg       | 1623 |
| Db | 1468 | GTCATCAATTCGGAAGGCTGGGAGAGTTTCACGTAACGCGCGGTGCTGTTCAATTT       | 1525 |
| Qy | 1624 | accattatgcagatggttggggaatttcaactgtaaacgaggggaggttctggtt        | 1681 |

RESULT 14

LOCUS A21895 5677 bp DNA PAT  
DEFINITION Nucleotide sequence of plasmid pMatL16.  
ACCESSION A21895  
NID 9641394

KEYWORDS  
SOURCE .  
unidentified.  
unidentified  
unclassified.

ORGANISM  
REFERENCE 1 (bases 1 to 5677)  
AUTHORS  
JOURNAL  
FEATURES

Patent: WO 9100353-A 4 10-JAN-1991;  
Location/Qualifiers  
1..5677

source  
CDS  
/organism="Artificial sequences"  
3882..5420  
/codon\_start=1  
/product="alph-amyase"  
/db\_xref="PID:9641395"

/translation="MKQOKRLVRLTLLELFLPHSAAANLNGTLMQFWMY  
PNDGHWKRLQNSAYLAHEGITAIPVAYKGTQADYGLDGLGFEHQGT  
RTKXGTGELQSAIKSLHSRDINVDVYINHGADATEDVAVEDPADNRVIG  
EHLIKATWHPFGSGTSDFKWHYHFDGTDWDSKRLNRYKFGKAWDEVSNE  
NGNDYDLYMADIDIDHPDAAEIKRWGTWYANLQDGLRDLAVKHIFSLRDWVNH  
VREKTGEMFTVAETWQNDGALENLKNFNHVSFVPLRYQFHAATOGGVDNR  
KLLNGTVVSKHPLKSVTFVNDHDTQPOQSLSTVQTFWFKPLAYAFILTRSGYPOVFE  
GDMTGTGDSQREIPALKHKIEPIKARKOYAYCAOHDFDHDHIDVIGWTRGDSVAN  
SGLAALITDGPGGAKRMVYGRNAGETWHDITGNRSPVINSSEGWEFHVNGSVSI  
YVQR"

BASE COUNT 1514 a 1315 c 1351 g 1497 t  
ORIGIN

Query Match 21.2%; Score 376; DB 17; Length 5677;  
Best Local Similarity 66.2%; Pred. No. 0.00e+00;  
Matches 952; Conservative 0; Mismatches 468; Indels 18; Gaps 9;



[illegible]

Job time : 2996 secs.

[illegible]